

ON SYMMETRIC LINEAR DIFFUSIONS

LIPING LI AND JIANGANG YING

ABSTRACT. The main purpose of this paper is to explore the structure of local and regular Dirichlet forms associated with symmetric linear diffusions. Let $(\mathcal{E}, \mathcal{F})$ be a regular and local Dirichlet form on $L^2(I, m)$, where I is an interval and m is a fully supported Radon measure on I . We shall first present a complete representation for $(\mathcal{E}, \mathcal{F})$, which shows that $(\mathcal{E}, \mathcal{F})$ lives on at most countable disjoint ‘effective’ intervals with corresponding scale function on each interval, and any point outside these intervals is a trap of the linear diffusion. Furthermore, we shall give a necessary and sufficient condition for $C_c^\infty(I)$ being a special standard core of $(\mathcal{E}, \mathcal{F})$ and identify the closure of $C_c^\infty(I)$ in $(\mathcal{E}, \mathcal{F})$ when $C_c^\infty(I)$ is contained but not necessarily dense in \mathcal{F} relative to the \mathcal{E}_1 -norm. This paper is partly motivated by a result of Hamza [9], stated in [8, Theorem 3.1.6] and provides a different point of view to this theorem. To illustrate our results, many examples are provided.

CONTENTS

1. Background and introduction	1
Notations	3
2. Representations of one-dimensional strongly local Dirichlet forms	4
2.1. Main result	4
2.2. Proof of Theorem 2.1	5
2.3. Effective intervals	12
2.4. When $I = \mathbb{R}$	14
3. When \mathcal{F} contains smooth functions	15
3.1. Regular Dirichlet subspaces	15
3.2. Irreducible case	16
3.3. General case	16
3.4. Recovering Hamza’s theorem	20
3.5. Closure of C_c^∞	22
4. Killing inside	25
Acknowledgement	26
References	26

1. BACKGROUND AND INTRODUCTION

Though the title of this paper is simple and easy to understand, it is still a little bit misleading because it actually focuses on the structure of regular and local Dirichlet forms on real line. However it is roughly correct in the sense that such Dirichlet forms are in one to one correspondence with symmetric linear diffusions. A Dirichlet form is a closed symmetric form with Markovian property on $L^2(E, m)$ space, where E is a nice topological space and m a Radon measure on E . Theory of Dirichlet form, established by a group of French mathematicians lead by Beurling, Choquet and Deny, is closely related to probability theory because of its Markovian property. We refer the notion and terminology in theory of Dirichlet forms to [1, 8, 14].

What do ‘regular’ and ‘local’ above mean? On one hand, the ‘regularity’ of a Dirichlet form assures that it is associated with a symmetric Markov process, due to a series of important works

2010 *Mathematics Subject Classification.* Primary 31C25, 60J55; Secondary 60J60.

Key words and phrases. Symmetric linear diffusions, Dirichlet forms, regular Dirichlet subspaces, Kolmogorov cycle test.

The first named author is partially supported by a joint grant (No. 2015LH0043) of China Postdoctoral Science Foundation and Chinese Academy of Science, and China Postdoctoral Science Foundation (No. 2016M590145). The second named author is partially supported by NSFC No. 11271240.

by M. Fukushima, M. L. Silverstein in 1970's, and Albeverio, Z. M. Ma, M. Röckner in 1990's. Note and be careful that the regularity of a Dirichlet form is different from the regularity of a linear diffusion. Intuitively the 'regularity' of a Dirichlet form means that if there are enough continuous functions in its domain, while the 'regularity' of a linear diffusion means roughly that it is irreducible as in (2.1). On the other hand, the locality of a Dirichlet form is equivalent to continuity of sample paths of the related Markov process. Hence a regular and local Dirichlet form on (or an interval of) real line corresponds to a linear diffusion, which is, by definition, a strong Markovian process with continuous paths on an interval of real line. One-dimensional Brownian motion is a typical example of linear diffusion. There are rich examples and interesting problems in linear diffusions, which have attracted many famous probabilists such as W. Feller, K. Itô and H. McKean, and the classical results are now collected in many books, such as [10, 11, 15]. Linear diffusion has simple and beautiful structure. It is well known that under the 'regularity' condition (see (2.1) below), a linear diffusion could be characterized essentially by a function, called scale function, a measure, called speed measure, and the killing measure if there is killing inside (see [15, Chapter V §7]). It was illustrated in [3] that in this case its speed measure is also the unique symmetric measure. Then regular linear diffusions are under the framework of symmetric Markov processes, which generate naturally Dirichlet forms, or energy forms, through infinitesimal generators of the transition semigroups. Therefore regular and local Dirichlet forms on real line may be viewed as symmetric linear diffusions.

As said above, the main purpose of this paper is to explore the structure of regular and local Dirichlet forms on real line. It is motivated by one of our recent paper [13] in which we felt that the structure might be characterized in some way. We would indicate that this paper is also motivated by an early result of Hamza, see [8, Theorem 3.1.6] and [9], which gives a precise characterization for a symmetric form defined on $C_c^\infty(\mathbb{R})$ to be closable. The proof of it is completely analytic. We are interested in and try to find its probabilistic and more natural explanation behind it.

To explore the structure, we are going to find a representation of such Dirichlet forms first. Representation theorem of any kind is always interesting and important, and lives everywhere in mathematics. For example, as mentioned above, a linear diffusion may be represented by its scale function and speed measure, and a Lévy process may be represented by its Lévy exponent, or more precisely by a non-negative definite matrix, Lévy measure and a drift vector. The most important representation theorem in the theory of Dirichlet forms is the famous Beurling-Deny decomposition, which asserts that a regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(E, m)$ may be decomposed into three parts: diffusion, jump, and killing, as follows,

$$\mathcal{E}(u, v) = \mathcal{E}^{(c)}(u, v) + \int_{E \times E \setminus d} (u(x) - u(y))(v(x) - v(y))J(dx dy) + \int_E u(x)v(x)k(dx),$$

for $u, v \in \mathcal{F} \cap C_c(E)$. As you see, though jump and killing parts may be characterized clearly by measures J and k respectively, the diffusion part $\mathcal{E}^{(c)}$ is given by a form with strongly local property and its representation is not completely clear. Therefore it is surely significant to find the representation of strongly local forms in various concrete cases. However such representations have seldom appeared in literatures. As we have seen, there are rich examples of regular Dirichlet forms in the classical textbook [8], including Example 1.2.2, corresponding to linear diffusion with natural scale, and Hamza's result, as stated in Theorem 3.1.6. However nothing about representation has been mentioned until two papers [3, 6], in which the authors gave the representation for a strongly local and irreducible regular Dirichlet form on \mathbb{R} independently, were published. In this paper, we shall first present a complete representation for a strongly local and regular Dirichlet form on \mathbb{R} , without assuming irreducibility. For a long time we have paid little attention to non-irreducible case because we believed that a non-irreducible process was simply made of some irreducible parts so that the generalization is trivial. As you will see in Theorem 2.1, the generalization from irreducible case to non-irreducible is far from routine. Especially, the proof of sufficiency, the regularity of such representation, needs more new ideas and techniques.

The advance in technique is a reason to start this paper, but it is not the most important reason. The main reason that we are interested in the Dirichlet forms without irreducibility is that they appear naturally in various cases so that the theory of symmetric linear diffusions would not be complete without including it. The Hamza's result is an example in which our

new representation theorem may be used to illustrate why his result was formulated this way probabilistically, and perfectly, as shown in Theorem 3.10. In the recent paper [13], we find that a regular Dirichlet extension of one-dimensional Brownian motion is associated to a linear diffusion, but not necessarily ‘regular’ or irreducible. It is proved in [13, Theorem 3.3] that every regular Dirichlet extension of one-dimensional Brownian motion may be essentially decomposed into at most countable disjoint invariant intervals and an exceptional set relative to this extension. To formulate the representation of strongly local regular Dirichlet forms on \mathbb{R} , we have to come back to those classical results for linear diffusions stated in K. Itô and H. Jr. McKean’s classical book [10, 11], which show us profound understanding about intuitive behaviours of general linear diffusion. Some of their results are essentially used in [13] and we shall also review several important lemmas in §2.2.

After the representation of strongly local regular Dirichlet forms is put forward, we shall treat a series of problems, related to the finer matters in structure, in which we are interested, for example, when the form takes $C_c^\infty(\mathbb{R})$ as a core and when the form only contains it. As we answer those problems, the Hamza’s result will be more clearly and intuitively shaped.

This paper is organized as follows. In §2, the main result (Theorem 2.1) derives a concrete formula to express the regular and strongly local Dirichlet form on an interval. It is shown that the state space can be decomposed into at most countable disjoint intervals, on each of which the linear diffusion is regular and characterized by a scale function. Furthermore, any point outside these intervals is a trap of the diffusion. In other words, the trajectory starting from the point outside the intervals will stay there forever. The special case of Theorem 2.1 on \mathbb{R} is given in Corollary 2.13 and a general version with killing inside of Theorem 2.1 is presented in Theorem 4.1.

A classical method to construct a regular Dirichlet space is to start from a closable and Markovian symmetric form with the domain C_c^∞ , see [8, §3]. Then the closure of C_c^∞ with respect to this form is a regular Dirichlet form. In §3, we shall explore the similar issue through a different angle, and give a necessary and sufficient condition for space of smooth functions, $C_c^\infty(I)$, to be a special standard core of (2.4), see Theorem 3.7. Applying Theorem 3.7 to Hamza’s result (see Theorem 3.10), we obtain a more intuitive description for structure of the form (3.10) and the associated symmetric linear diffusion, which is generated by the closure of $C_c^\infty(I)$, see Corollary 3.11. Moreover, we identify the closure of $C_c^\infty(I)$ in (2.4) when $C_c^\infty(I)$ is not necessarily a special standard core of (2.4), see Theorem 3.14. The possible killing insides of the associated symmetric linear diffusions are discussed in §4. Combining these results together, we think that the structure of strongly local regular Dirichlet forms on real line is almost completely clear.

Notations. Let us put some often used notations here for handy reference, though we may restate their definitions when they appear.

For $a < b$, $\langle a, b \rangle$ is an interval where a or b may or may not be contained in $\langle a, b \rangle$. Notations dx and $|\cdot|$ stand for the Lebesgue measure on \mathbb{R} or an interval throughout the paper if no confusion caused. The restrictions of a measure μ and a function f on I are denoted by $\mu|_I$ and $f|_I$ respectively. The notation ‘:=’ is read as ‘to be defined as’.

For a scale function \mathbf{s} (i.e. a continuous and strictly increasing function) on some interval J , $d\mathbf{s}$ represents its associated Lebesgue-Stieltjes measure on J . Set $\mathbf{s}(J) := \{\mathbf{s}(x) : x \in J\}$. For two measures μ and ν , $\mu \ll \nu$ means μ is absolutely continuous with respect to ν . Given a scale function \mathbf{s} on J and another function f on J , $f \ll \mathbf{s}$ means $f = g \circ \mathbf{s}$ for some absolutely continuous function g and $\frac{df}{d\mathbf{s}} := g' \circ \mathbf{s}$. Given an interval J , the classes $C_c(J)$, $C_c^1(J)$ and $C_c^\infty(J)$ denote the spaces of all continuous functions on J with compact support, all continuously differentiable functions with compact support and all infinitely differentiable functions with compact support, respectively.

Let $X = (X_t)_{t \geq 0}$ be a Markov process associated with a Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(E, m)$. If A is an invariant set of X (see [13, §2.1]), then the restriction of $(\mathcal{E}, \mathcal{F})$ to A is denoted by $(\mathcal{E}^A, \mathcal{F}^A)$ and the restriction of X to A is denoted by X^A . More precisely,

$$\mathcal{F}^A := \{f|_A : f \in \mathcal{F}\}, \quad \mathcal{E}^A(f|_A, g|_A) := \mathcal{E}(1_A f, 1_A g), \quad f, g \in \mathcal{F},$$

which is the associated Dirichlet form on $L^2(A, m|_A)$ of X^A . If U is an open subset of E , then the part Dirichlet form of $(\mathcal{E}, \mathcal{F})$ on U is denoted by $(\mathcal{E}_U, \mathcal{F}_U)$ and the part process of X on U is denoted by X_U .

2. REPRESENTATIONS OF ONE-DIMENSIONAL STRONGLY LOCAL DIRICHLET FORMS

This section is devoted to characterize the Dirichlet form associated to a symmetric linear diffusion. We are given an interval

$$I = \langle l, r \rangle,$$

where l or r may or may not be in I , and a fully supported Radon measure m on I . Let $(\mathcal{E}, \mathcal{F})$ be a regular and strongly local Dirichlet form on $L^2(I, m)$. Clearly, its associated Hunt process X is a symmetric diffusion process on I with no killing inside due to strong locality. The assumption of no killing inside has no essential influence on our results as we will explain later. A trap, by its name, is a point that X , starting from it, never leaves. Actually as illustrated later, the interior of I has no trap that X could reach from other points, and however the boundary could act like a trap that X could reach from other points, if I is open. See the examples in §2.3. Hence we need to pay extra attention to the boundaries of I . When X is irreducible (or ‘regular’ in the terminology of [15, (45.2)]) in the sense that

$$(2.1) \quad \mathbf{P}_x(\sigma_y < \infty) > 0, \quad \forall x, y \in I,$$

where σ_y is the hitting time of $\{y\}$ relative to X , its Dirichlet form $(\mathcal{E}, \mathcal{F})$ is characterized by its scale function as in [1, §2.2.3], [3] and [7]. However, the irreducibility is not always right, and we raised many reducible examples such as the regular Dirichlet extensions of one-dimensional Brownian motion in a recent study [13]. A natural and interesting question is how to describe this Dirichlet form without the irreducibility. This is what we shall do in this section.

2.1. Main result. Before presenting the main result, we need to prepare some terminologies and notations. Let $J := \langle a, b \rangle \subset I$ be an interval, where a or b may or may not be in J . Take a fixed point e in J as follows

$$e := \begin{cases} \frac{a+b}{2}, & |a| + |b| < \infty, \\ a+1, & a > -\infty, b = \infty, \\ b-1, & a = -\infty, b < \infty, \\ 0, & a = -\infty, b = \infty. \end{cases}$$

Denote all scale functions on J by $\mathbf{S}(J)$

$$\mathbf{S}(J) := \{\mathbf{s} : J \rightarrow \mathbb{R} : \mathbf{s} \text{ is continuous and strictly increasing, } \mathbf{s}(e) = 0\},$$

where we set $\mathbf{s}(e) = 0$ to assure the uniqueness. When a or b is not in J , we define $\mathbf{s}(a)$ or $\mathbf{s}(b)$ for $\mathbf{s} \in \mathbf{S}(J)$ to be the limit:

$$\mathbf{s}(a) := \lim_{x \downarrow a} \mathbf{s}(x) \geq -\infty, \quad \mathbf{s}(b) := \lim_{x \uparrow b} \mathbf{s}(x) \leq \infty.$$

We introduce the following conditions on a scale function $\mathbf{s} \in \mathbf{S}(J)$:

- (A) When $a > l$ or $a = l \in I$, $\mathbf{s}(a) = -\infty$ if and only if $a \notin J$;
- (B) When $b < r$ or $b = r \in I$, $\mathbf{s}(b) = \infty$ if and only if $b \notin J$.

What we need is the following subset of $\mathbf{S}(J)$:

$$(2.2) \quad \mathbf{S}_\infty(J) := \{\mathbf{s} \in \mathbf{S}(J) : \mathbf{s} \text{ satisfies (A) and (B)}\}.$$

For the Radon measure m on I , let $m_J := m|_J$ be its restriction to J . We write $m_J(\{b-\}) = \infty$ (resp. $< \infty$) to stand for that for some small constant $\epsilon > 0$, $m_J((b-\epsilon, b)) = \infty$ (resp. $< \infty$). The notation $m_J(\{a+\}) = \infty$ or $< \infty$ has the similar meaning. Note that $m_J(\{b-\}) = \infty$ (resp. $m_J(\{a+\}) = \infty$) might happen only when $b = r \notin I$ (resp. $a = l \notin I$) since m is Radon on I .

Next we introduce a Dirichlet form on $L^2(J, m_J)$ induced by a scale function $\mathbf{s} \in \mathbf{S}_\infty(J)$. The conditions (A) and (B) has no restriction on \mathbf{s} for the cases $a = l \notin I$ and $b = r \notin I$. Thus the following situations may happen:

- (L0) $a = l \notin I$, $\mathbf{s}(a) > -\infty$ and $m_J(a+) < \infty$;
- (R0) $b = r \notin I$, $\mathbf{s}(b) < \infty$ and $m_J(b-) < \infty$,

and in either case, we need to set up the Dirichlet boundary condition. Define a quadratic form $(\mathcal{E}^{(\mathbf{s})}, \mathcal{F}^{(\mathbf{s})})$ on $L^2(J, m_J)$ as follows:

$$(2.3) \quad \begin{aligned} \mathcal{F}^{(\mathbf{s})} &:= \left\{ u \in L^2(J, m_J) : u \ll \mathbf{s}, \frac{du}{d\mathbf{s}} \in L^2(J, d\mathbf{s}); \right. \\ &\quad \left. u(a) = 0 \text{ (resp. } u(b) = 0) \text{ whenever (L0) (resp. (R0)) holds} \right\}, \\ \mathcal{E}^{(\mathbf{s})}(u, v) &:= \frac{1}{2} \int_J \frac{du}{d\mathbf{s}} \frac{dv}{d\mathbf{s}} d\mathbf{s}, \quad u, v \in \mathcal{F}^{(\mathbf{s})}. \end{aligned}$$

Note that if $\mathbf{s}(a) > -\infty$ (resp. $\mathbf{s}(b) < \infty$), then any function u satisfying $u \ll \mathbf{s}$ and $du/d\mathbf{s} \in L^2(J, d\mathbf{s})$ admits the limit $u(a) := \lim_{x \downarrow a} u(x)$ (resp. $u(b) := \lim_{x \uparrow b} u(x)$) (Cf. [1, §2.2.3]). From [1, Theorem 2.2.11], it follows that $(\mathcal{E}^{(\mathbf{s})}, \mathcal{F}^{(\mathbf{s})})$ is a regular and strongly local Dirichlet form on $L^2(J, m_J)$ and its associated diffusion process $X^{(\mathbf{s})}$ is given by the scale function \mathbf{s} , the speed measure m_J and no killing inside. Moreover, when (L0) or (R0) happens, a or b is an absorbing boundary for $X^{(\mathbf{s})}$ and $X^{(\mathbf{s})}$ is also called a minimal diffusion on J in [1, Example 3.5.7]. For more discussions about the boundary behaviours of $X^{(\mathbf{s})}$, we refer to [1, §2.2.3], [7] and [13].

We are now in a position to state the main result of this section.

Theorem 2.1. *Let $I = \langle l, r \rangle$ be an interval as above and m a fully supported Radon measure on I . Then $(\mathcal{E}, \mathcal{F})$ is a regular and strongly local Dirichlet form on $L^2(I, m)$ if and only if there exist a set of at most countable disjoint intervals $\{I_n = \langle a_n, b_n \rangle : I_n \subset I, n \geq 1\}$ and a scale function $\mathbf{s}_n \in \mathbf{S}_\infty(I_n)$ for each $n \geq 1$ such that*

$$(2.4) \quad \begin{aligned} \mathcal{F} &= \left\{ u \in L^2(I, m) : u|_{I_n} \in \mathcal{F}^{(\mathbf{s}_n)}, \sum_{n \geq 1} \mathcal{E}^{(\mathbf{s}_n)}(u|_{I_n}, u|_{I_n}) < \infty \right\}, \\ \mathcal{E}(u, v) &= \sum_{n \geq 1} \mathcal{E}^{(\mathbf{s}_n)}(u|_{I_n}, v|_{I_n}), \quad u, v \in \mathcal{F}, \end{aligned}$$

where for each $n \geq 1$, $(\mathcal{E}^{(\mathbf{s}_n)}, \mathcal{F}^{(\mathbf{s}_n)})$ is given by (2.3) with the scale function \mathbf{s}_n on I_n . Moreover, the intervals $\{I_n : n \geq 1\}$ and scale functions $\{\mathbf{s}_n : n \geq 1\}$ are uniquely determined, if the difference of order is ignored.

Remark 2.2. We would like to give several remarks before the proof of this theorem.

- (1) A special case of Theorem 2.1 is considered in [13, Theorem 3.3], in which $(\mathcal{E}, \mathcal{F})$ is a so-called regular Dirichlet extension of one-dimensional Brownian motion. It turns out that in [13] $(\bigcup_{n \geq 1} I_n)^c$ has to be of zero Lebesgue measure (m is the Lebesgue measure in this case) and an additional condition on the scale function is imposed. For a general symmetric linear diffusion, these restrictions are not necessarily required.
- (2) Though the intervals are disjoint, they may have common endpoints. For example, $I_m = (a_m, b_m]$ and $I_n = (a_n, b_n)$ with $b_m = a_n$.
- (3) The set of intervals $\{I_n : n \geq 1\}$ is allowed to be empty. Particularly, $\mathcal{F} = L^2(I, m)$ and $\mathcal{E}(u, v) = 0$ for any $u, v \in \mathcal{F}$. Then X is a trivial diffusion process and every point of I is a trap of X .

2.2. Proof of Theorem 2.1. Recall that two m -symmetric Hunt processes are *equivalent* if they possess a common properly exceptional set outside which their transition functions coincide (Cf. [8, Theorem 4.2.8]). Any regular and strongly local Dirichlet form on $L^2(I, m)$ is associated with a class of equivalent m -symmetric diffusions, in which there is a representative X satisfying

$$(2.5) \quad \mathbf{P}_x(X_{\zeta-} \in I, \zeta < \infty) = 0$$

for every $x \in I$, where ζ is the lifetime of X . This property is usually called ‘no killing inside’.

We shall use the same notations as in [13, §3.2] to classify the state space I of the linear diffusion X as follows (Cf. also [11]):

- (1) Λ_2 : regular points, which could reach both sides,
- (2) $\Lambda_r \cup \Lambda_l$: singular points, which could reach at most one side,
- (3) Λ_l / Λ_r : left/right singular points, which could not reach right/left side,
- (4) $\Lambda_{pl} / \Lambda_{pr}$: left/right shunt points, which could reach only left/right side,
- (5) Λ_t : traps, the points which could not reach either side.

Keep in mind the properties of these classes are presented in [13, Lemma 3.7]. Especially, we can deduce the following lemma under the symmetry as in [13, Lemmas 3.8 and 3.9].

Lemma 2.3. (1) Fix a right (resp. left) singular point $b \in \Lambda_r$ (resp. Λ_l). Under the symmetry, for any $a < b$ (resp. $a > b$), it holds that

$$\mathbf{P}_a(\sigma_b < \infty) = 0,$$

where σ_b is hitting time of $\{b\}$ relative to the diffusion process. Particularly, under the symmetry if $b \in \Lambda_t$ then for any $a \neq b$,

$$\mathbf{P}_a(\sigma_b < \infty) = 0.$$

(2) Under the symmetry, if $(a, b) \subset \Lambda_r$ (resp. Λ_l), then $(a, b) \subset \Lambda_t$.

Remark 2.4. Although the discussions in [13] are made on the one-dimensional Euclidean space \mathbb{R} , all the results in [13, §3.3.1-3.3.2] still hold for the diffusion on I with no killing inside. For the classification of the points, every point x in the interior of I (i.e. $x \in (l, r)$) can be classified as [13, Definition 3.6]. For the possible $x = r \in I$, formally we say x is a left singular point, i.e. $x \in \Lambda_l$, and we have

$$e^- := \mathbf{P}_x(\sigma_{x-} = 0) = 0 \text{ or } 1,$$

where $\sigma_{x-} := \inf\{t > 0 : X_t < x\}$, by the Blumenthal 0-1 law. We say x is a left shunt point, i.e. $x \in \Lambda_{pl}$ if $e^- = 1$ and a trap, i.e. $x \in \Lambda_t$ if $e^- = 0$. Particularly, if $x = r$ is a trap, then

$$\mathbf{P}_x(X_t = x, \forall t \geq 0) = 1$$

due to ‘no killing inside’. Similarly, we can classify the possible $x = l \in I$ as a right shunt point or a trap.

2.2.1. Proof of necessity. Let $(\mathcal{E}, \mathcal{F})$ be a regular and strongly local Dirichlet form on $L^2(I, m)$ and X a representative of the associated m -symmetric diffusions satisfying (2.5) for every $x \in I$.

First note that the regular set Λ_2 is open (Cf. [13, Lemma 3.7 (6)]). Thus we may write Λ_2 as a union of disjoint open intervals:

$$(2.6) \quad \Lambda_2 = \bigcup_{n \geq 1} (a_n, b_n).$$

Let

$$F := \Lambda_2^c = \Lambda_l \cup \Lambda_r,$$

which is closed. Further let $\overset{\circ}{F}$ be the interior of F and $\partial F := F \setminus \overset{\circ}{F}$. Clearly, $\{a_n, b_n : n \geq 1\} \cap I \subset \partial F$, ∂F is nowhere dense and $\overset{\circ}{F}$ is open. The proof of necessity will be completed in four steps.

Step 1: Every point in $\overset{\circ}{F}$ is a trap.

Lemma 2.5. $\overset{\circ}{F} \subset \Lambda_t$.

Proof. Take $x \in (c, d) \subset \overset{\circ}{F}$. Suppose that $x \in \Lambda_{pr}$. Then there exists a small constant $0 < \epsilon < |d - x|$ such that $(x, x + \epsilon) \subset \Lambda_r$. In fact, if this assertion is not right, then there is a decreasing sequence $\{x_k : k \geq 1\} \subset \Lambda_{pl}$ in (x, d) satisfying $x_k \downarrow x$ as $k \rightarrow \infty$. By [13, Lemma 3.7 (5)], we have $x \in \Lambda_l$, which contradicts the fact $x \in \Lambda_{pr}$.

Since $(x, x + \epsilon) \subset \Lambda_r$, it follows from Lemma 2.3 (2) that $(x, x + \epsilon) \subset \Lambda_t$. But $x \in \Lambda_{pr}$ implies that for some $y > x$ (Cf. [13, Lemma 3.7 (4)])

$$\mathbf{P}_x(\sigma_y < \infty) > 0.$$

Clearly, y can be taken in $(x, x + \epsilon)$. This contradicts Lemma 2.3 (1). Therefore, we have $x \notin \Lambda_{pr}$. Similarly we can also deduce that $x \notin \Lambda_{pl}$. Finally we conclude $x \in \Lambda_t$. That completes the proof. \square

Step 2: We treat the points in ∂F .

Since $\overset{\circ}{F}$ is open, we also write it as a union of disjoint open intervals:

$$(2.7) \quad \overset{\circ}{F} = \bigcup_{m \geq 1} (c_m, d_m).$$

Then ∂F may be divided into three families: $\{a_n, b_n : n \geq 1\} \cap I$, $\{c_m, d_m : m \geq 1\} \cap \partial F$ and others. We denote them by $\partial_1 F$, $\partial_2 F$ and $\partial_3 F$ respectively.

Consider a point $b_n \in \partial_1 F$ for some n . There are the following possible situations:

- (1) $b_n \in \Lambda_t$. Notice the first assertion in Lemma 2.3.
- (2) $b_n \in \Lambda_{pl}$. This implies that $\mathbf{P}_{b_n}(\sigma_y < \infty) > 0$ for any $y \in (a_n, b_n)$ (Cf. Lemma 3.7 (4) of [13]), but $\mathbf{P}_{b_n}(\sigma_z < \infty) = \mathbf{P}_z(\sigma_{b_n} < \infty) = 0$ for any $z > b_n$ (Cf. [13, Lemma 3.7 (2)] and Lemma 2.3 (1)). Hence if $\mathbf{P}_y(\sigma_{b_n} < \infty) > 0$ for some (equivalently, all) $y \in (a_n, b_n)$, then we have

$$(2.8) \quad \mathbf{P}_y(\sigma_{b_n} < \infty) \cdot \mathbf{P}_{b_n}(\sigma_y < \infty) > 0, \quad \forall y \in (a_n, b_n).$$

Otherwise, $m(\{b_n\}) = 0$ and $\{b_n\}$ is an m -polar set relative to X by Remark 2.6 below.

- (3) $b_n \in \Lambda_{pr}$. Clearly, for any $y \in (a_n, b_n)$, $\mathbf{P}_y(\sigma_{b_n} < \infty) = \mathbf{P}_{b_n}(\sigma_y < \infty) = 0$. Moreover, by Remark 2.7 below, we know that $b_n = a_j$ for some j . In addition, similar to Remark 2.6, either it holds

$$(2.9) \quad \mathbf{P}_y(\sigma_{a_j} < \infty) \cdot \mathbf{P}_{a_j}(\sigma_y < \infty) > 0, \quad \forall y \in (a_j, b_j)$$

or $\{b_n\}$ is an m -polar set relative to X .

The point $a_n \in \partial F_1$ can be treated similarly as above.

Remark 2.6. Let us prove the following fact: if $b_n \in \Lambda_{pl}$ and $\mathbf{P}_y(\sigma_{b_n} < \infty) = 0$ for all $y \in (a_n, b_n)$, then $m(\{b_n\}) = 0$ and thus $\{b_n\}$ is m -polar relative to X . Suppose that $m(\{b_n\}) > 0$. Recall that the probability transition semigroup $(P_t)_{t \geq 1}$ satisfies the following symmetric property:

$$(2.10) \quad (P_t f, g)_m = (f, P_t g)_m, \quad f, g \in \mathcal{B}_+(I), t \geq 0,$$

where $(f, g)_m$ stands for the inner product of $L^2(I, m)$ and $\mathcal{B}_+(I)$ is the set of all non-negative Borel measurable functions on I . Take $f(x) := 1_{\{b_n\}}(x)$ and

$$g(x) := \begin{cases} -x + b_n, & x \leq b_n, \\ 0, & x > b_n. \end{cases}$$

Since for any $x \neq b_n$,

$$P_t f(x) = \mathbf{P}_x(X_t = b_n) = 0$$

and $g(b_n) = 0$, the left side of (2.10) equals 0. It follows from $m(\{b_n\}) > 0$ that $P_t g(b_n) = 0$, which implies $\mathbf{P}_{b_n}(X_t < b_n) = 0$ for any $t \geq 0$. It follows that

$$\mathbf{P}_{b_n} \left(\bigcup_{t \in \mathbb{Q}_+} \{X_t < b_n\} \right) = 0$$

and

$$1 = \mathbf{P}_{b_n} \left(\bigcap_{t \in \mathbb{Q}_+} \{X_t \geq b_n\} \right) = \mathbf{P}_{b_n} \left(\bigcap_{t \geq 0} \{X_t \geq b_n\} \right).$$

This indicates $b_n \in \Lambda_r$, which contradicts the fact $b_n \in \Lambda_{pl}$.

Remark 2.7. For the case $b_n \in \Lambda_{pr}$, we show that $b_n = a_j$ for some j . If not, then either $b_n = c_k$ (see (2.7)) for some k or a subsequence $\{(a_{j'}, b_{j'})\}$ of (2.6) decreases to b_n . Both of them are impossible. For the former one, it follows from Lemma 2.5 that $(c_k, d_k) \subset \Lambda_t$. Then $\mathbf{P}_{b_n}(\sigma_y < \infty) = 0$ for any $y > b_n$ by Lemma 2.3 (1). This contradicts that $b_n \in \Lambda_{pr}$. The latter one also contradicts that $b_n \in \Lambda_{pr}$ since it implies that $\mathbf{P}_{b_n}(\sigma_y < \infty) > 0$ for some $y > b_n$.

Next, consider $d_m \in \partial_2 F$ (the case $c_m \in \partial_2 F$ is similar). Clearly $d_m \in \Lambda_r$ since $(c_m, d_m) \subset \Lambda_t$. Moreover, either $d_m = a_j$ for some j or a subsequence $\{(a_{j'}, b_{j'})\}$ of (2.6) decreases to d_m . Similar to the third case of b_n above, we can deduce that either of the followings may happen:

- (1) $d_m \in \Lambda_t$.
- (2) $d_m \in \Lambda_{pr}$. Then it must hold that $d_m = a_j$ for some j . In addition, either it holds

$$\mathbf{P}_y(\sigma_{a_j} < \infty) \cdot \mathbf{P}_{a_j}(\sigma_y < \infty) > 0, \quad \forall y \in (a_j, b_j)$$

or $\{d_m\}$ is m -polar relative to X .

Finally, we assert that $\partial_3 F \subset \Lambda_t$. In fact, for any $x \in \partial_3 F$, there exist two subsequences of $\{(a_n, b_n) : n \geq 1\} \cup \{(c_m, d_m) : m \geq 1\}$ such that one decreases to x and the other increases to x . This excludes the possibility of $x \in \Lambda_{pr} \cup \Lambda_{pl}$. Therefore $\partial F_3 \subset \Lambda_t$.

In a word, except for an m -polar set, the point in F either belongs to Λ_t , or is the endpoint of some (a_n, b_n) with the property (2.8) or (2.9). Furthermore, the excluded m -polar set is contained in $\{a_n, b_n : n \geq 1\} \cap I$.

Step 3: Irreducible intervals and the diffusions on them.

Add the endpoints of every (a_n, b_n) with the property (2.8) or (2.9) into (a_n, b_n) and we obtain a new interval $I_n := \langle a_n, b_n \rangle$. First note that $I_n \cap I_m = \emptyset$ for $n \neq m$. In fact, $I_n \cap I_m \neq \emptyset$ only occurs when $a_n = b_m$ or $a_m = b_n$. This concludes that either $a_n = b_m \in \Lambda_2$ or $a_m = b_n \in \Lambda_2$, which contradicts that (2.6). We shall characterize the restriction of X to I_n .

Fix n and let us consider the interval I_n . From the second step, we know that

$$\begin{aligned} \mathbf{P}_x(\sigma_y < \infty) &= 0, \quad \forall x \in I_n, y \in I \setminus I_n, \\ \mathbf{P}_y(\sigma_x < \infty) &= 0, \quad \forall x \in I_n, \text{ q.e. } y \in I \setminus I_n, \end{aligned}$$

where the excluded possible m -polar set is contained in $\{a_n, b_n\}$. Particularly, I_n is an invariant set of X (Cf. [13, §3.1]). Thus the restriction X^{I_n} of X to I_n is still a diffusion process on I_n and symmetric with respect to $m_n := m|_{I_n}$ with no killing inside. Furthermore, X^{I_n} is irreducible in the sense that

$$\mathbf{P}_x(\sigma_y < \infty) > 0, \quad \forall x, y \in I_n.$$

It follows that X^{I_n} is characterized uniquely by a scale function $\mathbf{s}_n \in \mathbf{S}(I_n)$ since the speed measure m_n is fixed (Cf. [3, §3]). We assert that $\mathbf{s}_n \in \mathbf{S}_\infty(I_n)$. For the condition **(A)**, if $a_n > l$, then $a_n \notin I_n$ implies $\mathbf{P}_x(\sigma_{a_n} < \infty) = 0$ for any $x \in I_n$, which is equivalent to (Cf. [1, (3.5.13)])

$$\int_{a_n}^c m_n((x, c)) d\mathbf{s}_n(x) = \infty$$

for $c \in (a_n, b_n)$. Note that $m_n((a_n, c)) < \infty$. Thus $\mathbf{s}_n(a_n) = -\infty$. Similarly we can also prove that $a_n \in I_n$ implies $\mathbf{s}_n(a_n) > -\infty$. For the case $a_n = l \in I$, we need only to note that

$$0 < m_n((a_n, c)) \leq m([l, c]) < \infty.$$

The condition **(B)** can be verified similarly. Therefore, it follows from the discussions above and [3, Theorem 3.1] (see also [1, §2.2.3 and Theorem 3.5.8]) that the associated Dirichlet form of X^{I_n} is given by (2.3) with \mathbf{s}_n, I_n instead of \mathbf{s}, J respectively.

Step 4: The expression (2.4) of $(\mathcal{E}, \mathcal{F})$ can be deduced by [13, Corollary 3.11] and the following fact: $P_t f(x) = f(x)$ for $x \in \Lambda_t$. That completes the proof of necessity.

2.2.2. Proof of sufficiency. Let $(\mathcal{E}, \mathcal{F})$ be given by (2.4). Similar to [13, §3.2.5], we can deduce that $(\mathcal{E}, \mathcal{F})$ is a Dirichlet form in the wide sense on $L^2(I, m)$ by [13, Theorem 3.10]. It suffices to prove the regularity. We shall do it in two steps.

First step: To prove that $\mathcal{F} \cap C_c(I)$ is dense in $C_c(I)$ relative to the uniform norm.

For each n , take another scale function $\tilde{\mathbf{s}}_n \in \mathbf{S}(I_n)$ such that

$$\begin{aligned} d\tilde{\mathbf{s}}_n &\ll d\mathbf{s}_n, \\ \frac{d\tilde{\mathbf{s}}_n}{d\mathbf{s}_n} &= 0 \text{ or } 1, \text{ } d\mathbf{s}_n\text{-a.e.}, \\ |\tilde{\mathbf{s}}_n(b_n) - \tilde{\mathbf{s}}_n(a_n)| &\leq |b_n - a_n|. \end{aligned}$$

For the existence of $\tilde{\mathbf{s}}_n$, we refer to [13, Remark 3.1]. Define a measure λ on I as follows:

$$\lambda := \sum_{n \geq 1} d\tilde{\mathbf{s}}_n + dx|_{(\cup_{n \geq 1} I_n)^c}.$$

One may easily check that λ is a Radon measure on I and we denote its associated increasing function by \mathbf{s} , i.e. $\lambda = d\mathbf{s}$. Without loss of generality, we impose $\mathbf{s}(e) = 0$ where e is a fixed point in I . We assert that $\mathbf{s} \in \mathbf{S}(I)$. In fact, since $\lambda(\{x\}) = 0$ for any $x \in I$, \mathbf{s} is continuous. On the other hand, $\lambda((c, d)) > 0$ for any $c, d \in I$ with $c < d$ by $\tilde{\mathbf{s}}_n \in \mathbf{S}(I_n)$. This implies that \mathbf{s} is strictly increasing.

Consider now the following class of continuous functions:

$$\mathcal{C} := \{\varphi \circ \mathbf{s} : \varphi \in C_c^\infty(J)\},$$

where $J := \mathbf{s}(I) = \langle \mathbf{s}(a), \mathbf{s}(b) \rangle$ and $\mathbf{s}(a) \in J$ (resp. $\mathbf{s}(b) \in J$) if and only if $a \in I$ (resp. $b \in I$). Since \mathbf{s} is strictly increasing, it follows from Stone-Weierstrass theorem that \mathcal{C} is dense in $C_c(I)$ relative to the uniform norm. We need only to prove $\mathcal{C} \subset \mathcal{F} \cap C_c(I)$. Take a function $u = \varphi \circ \mathbf{s}$. Clearly, $u \in L^2(I, m)$. For each n , $u|_{I_n} = \varphi \circ \mathbf{s}|_{I_n} \ll \mathbf{s}_n$ since $d\mathbf{s}|_{I_n} = d\tilde{\mathbf{s}}_n \ll d\mathbf{s}_n$. Moreover,

$$\begin{aligned} \sum_{n \geq 1} \int_{I_n} \left(\frac{du|_{I_n}}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n &= \sum_{n \geq 1} \int_{I_n} \left(\frac{du|_{I_n}}{d\tilde{\mathbf{s}}_n} \right)^2 d\tilde{\mathbf{s}}_n \\ &= \sum_{n \geq 1} \int_{I_n} (\varphi' \circ \mathbf{s})^2 d\tilde{\mathbf{s}}_n \\ &\leq \int_I (\varphi' \circ \mathbf{s})^2 d\mathbf{s} \\ &< \infty. \end{aligned}$$

For the possible situations (L0) and (R0), clearly we have $u(a) = 0$ and $u(b) = 0$ respectively. Hence we conclude that $u \in \mathcal{F}$.

Second step: To prove that $\mathcal{F} \cap C_c(I)$ is dense in \mathcal{F} relative to the norm $\|\cdot\|_{\mathcal{E}_1}$.

Take $u \in \mathcal{F}$ and set

$$u_1 := u \cdot 1_{\cup_{n \geq 1} I_n}, \quad u_2 := u - u_1.$$

One may see $u_1, u_2 \in \mathcal{F}$ and $\mathcal{E}(u_2, u_2) = 0$. Fix a small constant $\epsilon > 0$. What we shall do is to find a function $u_\epsilon \in \mathcal{F} \cap C_c(I)$ such that $\mathcal{E}_1(u_\epsilon - u, u_\epsilon - u) < C \cdot \epsilon$ for some constant C .

We first treat u_1 . Since $(\mathcal{E}^{(\mathbf{s}_n)}, \mathcal{F}^{(\mathbf{s}_n)})$ is regular on $L^2(I_n, m_n)$ and $\mathcal{E}_1(u_1, u_1) < \infty$, we can take an integer N large enough and for each $1 \leq n \leq N$ a function $g_n \in \mathcal{F}^{(\mathbf{s}_n)} \cap C_c(I_n)$ such that

$$\begin{aligned} \sum_{n > N} \mathcal{E}_1^{(\mathbf{s}_n)}(u|_{I_n}, u|_{I_n}) &< \epsilon, \\ \mathcal{E}_1^{(\mathbf{s}_n)}(u|_{I_n} - g_n, u|_{I_n} - g_n) &< \epsilon/N, \quad 1 \leq n \leq N. \end{aligned}$$

Define a function $g \in \mathcal{F}$ as follows: $g|_{I_n} := u|_{I_n}$ for $1 \leq n \leq N$ and $g := 0$ elsewhere. Then g is continuous on I_n for $1 \leq n \leq N$ and

$$\mathcal{E}_1(u_1 - g, u_1 - g) < 2\epsilon.$$

Mimicking the proof of [13, Lemma 3.14] and referring to Remark 2.8 below, we can construct a function $g_\epsilon \in \mathcal{F} \cap C_c(I)$ such that $\mathcal{E}_1(g_\epsilon - g, g_\epsilon - g) < \epsilon$. This implies

$$\mathcal{E}_1(g_\epsilon - u_1, g_\epsilon - u_1) < 6\epsilon.$$

Remark 2.8. Similar to [13, Lemma 3.14], the discontinuous points of g are contained in the set of endpoints $\{a_n, b_n : 1 \leq n \leq N\} \setminus \{l, r\}$. Consider the right discontinuous endpoint $c := b_n$, we can adjust g through the method in [13, Lemma 3.14] except for one possible additional case: $c \in I_n, g(c) \neq 0$ and $(c, c + \delta) \subset (\cup_{n \geq 1} I_n)^c$ for some $\delta > 0$. Note that in the context of [13], this is impossible since $m((\cup_{n \geq 1} I_n)^c) = 0$ by [13, Theorem 3.3]. However, we do not have this restriction in this theorem. This additional case is easy to solve. Without loss of generality, assume that $g(c) = 1$. Notice that $g \equiv 0$ on $(c, c + \delta)$. Take a constant $\delta' < \delta$ with $m((c, c + \delta')) < \epsilon$ and set

$$\phi(x) := \begin{cases} 1 - (1/\delta') \cdot (x - c), & c \leq x \leq c + \delta' \\ 0, & \text{elsewhere.} \end{cases}$$

Then $g + \phi$ is continuous at c and

$$\mathcal{E}_1(\phi, \phi) = \int_I \phi^2 dm < \epsilon.$$

Next we treat u_2 . Since $u_2 \in L^2(I, m)$, it follows from [1, Exercise 2.13] that there exists a function $h \in \mathcal{F} \cap C_c(I)$ such that $\|h - u_2\|_{L^2(I, m)}^2 < \epsilon$. The key step to complete the proof is to construct a function h_ϵ through h so that it satisfies not only

$$h_\epsilon \in \mathcal{F} \cap C_c(I), \quad \|h_\epsilon - u_2\|_{L^2(I, m)}^2 < \tilde{C}\epsilon$$

but also $\mathcal{E}(h_\epsilon, h_\epsilon) < \tilde{C}\epsilon$ with some constant \tilde{C} . The construction of h_ϵ is given in the following lemma.

Lemma 2.9. Assume $u_2 \in L^2(I, m)$ with $u_2|_{\cup_{n \geq 1} I_n} = 0$ and fix a small constant $\epsilon > 0$. Then there exists a function $h_\epsilon \in \mathcal{F} \cap C_c(I)$ such that

$$\|h_\epsilon - u_2\|_{L^2(I, m)}^2 < 18\epsilon, \quad \mathcal{E}(h_\epsilon, h_\epsilon) < 6\epsilon.$$

Proof. Let h be above, i.e. $h \in \mathcal{F} \cap C_c(I)$ such that $\|h - u_2\|_{L^2(I, m)}^2 < \epsilon$. Since

$$\mathcal{E}(h, h) = \sum_{n \geq 1} \mathcal{E}^{(s_n)}(h|_{I_n}, h|_{I_n}) < \infty,$$

there exists an integer M such that $\sum_{n > M} \mathcal{E}^{(s_n)}(h|_{I_n}, h|_{I_n}) < \epsilon$. Recall that e_n is a fixed point in (a_n, b_n) . To attain h_ϵ , we shall change the values of h near the intervals $\{I_n : 1 \leq n \leq M\}$.

We first classify the right endpoints $\{b_n : 1 \leq n \leq M\}$ as follows:

- $b_n \in \mathcal{R}_1$ if $b_n = a_j$ for some $1 \leq j \leq M$;
- $b_n \in \mathcal{R}_2$ if $b_n \notin \mathcal{R}_1$ and $b_n \notin I_n$;
- $b_n \in \mathcal{R}_3$ if $b_n \notin \mathcal{R}_1$ and $b_n \in I_n$.

Similarly we can classify the left endpoints $\{a_n : 1 \leq n \leq M\}$ as $\mathcal{L}_1, \mathcal{L}_2$ and \mathcal{L}_3 . For each $1 \leq n \leq M$, we shall take a small constant $\varepsilon_n \geq 0$ and define a continuous function h_n^+ on $[e_n, b_n + \varepsilon_n]$ with $h_n^+(e_n) = 0$. The construction is as follows.

(1) $b_n \in \mathcal{R}_1$. Take $\varepsilon_n = 0$.

(1a) $b_n \in I_n$ or $a_j \in I_j$: define $h_n^+(x) = 0$ for any $x \in [e_n, b_n]$.

(1b) $b_n \notin I_n$ and $a_j \notin I_j$. In this case, $s_n(b_n) = \infty$. Take a constant $\delta > 0$. Mimicking (1a) of [13, Lemma 3.14], we can take another constant $0 < \delta' < \delta$ and define h_n^+ on $[e_n, b_n]$ as follows:

$$(2.11) \quad h_n^+(x) := \begin{cases} 0, & e_n \leq x \leq b_n - \delta, \\ h(b_n) \cdot \frac{s_n(x) - s_n(b_n - \delta)}{s_n(b_n - \delta') - s_n(b_n - \delta)}, & b_n - \delta < x < b_n - \delta', \\ h(b_n), & b_n - \delta' \leq x \leq b_n. \end{cases}$$

Regard h_n^+ as a function on I_n by letting $h_n^+(x) = 0$ on $\langle a_n, e_n \rangle$. Then clearly $h_n^+ \in \mathcal{F}^{(s_n)}$ and let $\delta > 0$ be small enough and then $\delta' < \delta$ be small enough, we have

$$\begin{aligned} \mathcal{E}_1^{(s_n)}(h_n^+, h_n^+) &\leq h(b_n)^2 \cdot m((b_n - \delta, b_n)) + \frac{1}{2} h(b_n)^2 \frac{1}{s_n(b_n - \delta') - s_n(b_n - \delta)} \\ &< \frac{\epsilon}{M}. \end{aligned}$$

(2) $b_n \in \mathcal{R}_2$. Take $\varepsilon_n = 0$.

(2a) $b_n \notin I$, in other words, $b_n = r \notin I$: define $h_n^+(x) = 0$ for any $x \in [e_n, b_n]$.

(2b) $b_n \in I$: define h_n^+ as (2.11) and we have $\mathcal{E}_1^{(s_n)}(h_n^+, h_n^+) < \epsilon/M$.

(3) $b_n \in \mathcal{R}_3$ and $b_n \neq r$.

(3a) $b_n = a_k$ for some $k > M$. Take $\varepsilon_n = b_k - b_n$.

In this case, $a_k \notin I_k$ and $s_k(a_k) = -\infty$. Let $h_n^+(x) := 0$ for any $x \in [e_n, b_n]$ and mimicking the case (1b), take $0 < \delta' < \delta$ small enough and define

$$(2.12) \quad h_n^+(x) := \begin{cases} 0, & a_k \leq x \leq a_k + \delta', \\ h(a_k + \delta) \cdot \frac{s_k(x) - s_k(a_k + \delta')}{s_n(a_k + \delta) - s_n(a_k + \delta')}, & a_k + \delta' < x < a_k + \delta, \\ h(x), & a_k + \delta \leq x < b_k. \end{cases}$$

When δ, δ' are small enough, we have

$$(2.13) \quad \mathcal{E}_1^{(s_k)}((h_n^+ - h)|_{I_k}, (h_n^+ - h)|_{I_k}) < \frac{\epsilon}{M}.$$

(3b) For any $\varepsilon > 0$, $(b_n, b_n + \varepsilon)$ contains at least one non-closed interval I_k with $k > M$.

In this case, without loss of generality, assume

$$(2.14) \quad \int_{(b_n, b_n + \varepsilon)} h(x)^2 m(dx) < \epsilon/M,$$

and $I_k = (a_k, b_k) \subset (b_n, b_n + \varepsilon)$ with $k > M$.

Take $\varepsilon_n = b_k - b_n$. Define $h_n^+(x) := 0$ for any $x \in [e_n, a_k)$ and for $x \in [a_k, b_k)$, define $h_n^+(x)$ as (2.12). Similarly, (2.13) holds.

(3c) For any $\varepsilon > 0$, $(b_n, b_n + \varepsilon)$ contains at least one interval I_j with $j > M$ and all of them are closed. Without loss of generality, assume $b_n + \varepsilon \notin \bigcup_{j \geq 1} \overset{\circ}{I}_j$ and ε is small enough so that (2.14) holds.

Take $\varepsilon_n = \varepsilon$ and set $h_n^+(x) = 0$ for $x \in [e_n, b_n]$. By mimicking the case (2) of [13, Lemma 3.14] and [13, Remark 3.15] we can construct a Cantor-type function h_n^+ on $[b_n, b_n + \varepsilon_n]$ such that $h_n^+(b_n) = 0$, $h_n^+(b_n + \varepsilon_n) = h(b_n + \varepsilon_n)$ and h_n^+ is a constant on every interval I_j , which is contained in $(b_n, b_n + \varepsilon)$. (Although $\bigcup_{n \geq 1} I_n$ is not necessarily dense in I , this construction is not difficult to realize by small modifications of [13, Remark 3.15]). Without loss of generality, we also have

$$(2.15) \quad \int_{(b_n, b_n + \varepsilon_n)} h_n^+(x)^2 m(dx) < \frac{\varepsilon}{M}.$$

(3d) For some $\varepsilon > 0$, $(b_n, b_n + \varepsilon) \subset \left(\bigcup_{n \geq 1} I_n\right)^c$.

Take $\varepsilon_n < \varepsilon$ small enough and set

$$h_n^+(x) := \begin{cases} 0, & x \in [e_n, b_n], \\ h(b_n + \varepsilon_n) \cdot \frac{x - b_n}{\varepsilon_n}, & b_n < x < b_n + \varepsilon_n \end{cases}$$

Without loss of generality, we have (2.14) and (2.15) hold.

(4) $b_n \in \mathcal{R}_3$ and $b_n = r$.

Take $\varepsilon_n = 0$ and define $h_n^+(x) := 0$ for $x \in [e_n, b_n]$.

Mimicking the procedures above, we can also take another small constant $\varepsilon'_n \geq 0$ and define a continuous function h_n^- on $(a_n - \varepsilon'_n, e_n]$ with $h_n^-(e_n) = 0$. Now we can define the function h_ε as follows:

$$h_\varepsilon(x) := \begin{cases} h_n^+(x), & x \in [e_n, b_n + \varepsilon_n), 1 \leq n \leq M, \\ h_n^-(x), & x \in (a_n - \varepsilon'_n, e_n], 1 \leq n \leq M, \\ 0, & x = b_n = a_j \text{ for case (1a)}, \\ h(x), & \text{others.} \end{cases}$$

From the construction of h_n^\pm , we can easily see that $h_\varepsilon \in C_c(I)$.

For any function g on I such that $2\mathcal{E}(g, g) = \sum_{n \geq 1} \int_{I_n} (dg/ds_n)^2 ds_n < \infty$ and any subset $J \subset I$, we regard $g|_J$ as a function on I by letting $g|_J(x) = 0$ for $x \in I \setminus J$ and set

$$\mathcal{E}(g|_J, g|_J) := \frac{1}{2} \sum_{n \geq 1} \int_{I_n \cap J} \left(\frac{dg}{ds_n} \right)^2 ds_n.$$

Let A be the set of all the endpoints $\{b_n : 1 \leq n \leq M\}$ in case (1a) and set

$$G := \left(\bigcup_{n=1}^M (a_n - \varepsilon'_n, b_n + \varepsilon_n) \right)^c \setminus A.$$

Then we have

$$\mathcal{E}(h_\varepsilon, h_\varepsilon) = \sum_{n=1}^M \mathcal{E}(h_n^+, h_n^+) + \sum_{n=1}^M \mathcal{E}(h_n^-, h_n^-) + \mathcal{E}(h|_G, h|_G).$$

For the cases (1), (2), (3c), (3d) and (4), clearly $\mathcal{E}(h_n^+, h_n^+) < \varepsilon/M$ and for the cases (3a) and (3b), it follows from (2.13) that

$$\mathcal{E}(h_n^+, h_n^+) \leq \frac{2\varepsilon}{M} + 2\mathcal{E}(h|_{I_k}, h|_{I_k}),$$

where I_k is the interval given in (3a) and (3b). The similar estimate holds for $\mathcal{E}(h_n^-, h_n^-)$. This implies $h_\varepsilon \in \mathcal{F}$ and

$$\mathcal{E}(h_\varepsilon, h_\varepsilon) \leq \frac{2\varepsilon}{M} \cdot M + \frac{2\varepsilon}{M} \cdot M + 2 \sum_{k > M} \mathcal{E}^{(s_k)}(h|_{I_k}, h|_{I_k}) < 6\varepsilon.$$

On the other hand,

$$\begin{aligned} & \int_I |h_\epsilon - u_2|^2 dm \\ &= \sum_{n=1}^M \int_{[e_n, b_n + \epsilon_n)} (h_n^+ - u_2)^2 dm + \sum_{n=1}^M \int_{(a_n - \epsilon'_n, e_n]} (h_n^- - u_2)^2 dm + \int_G (h - u_2)^2 dm \end{aligned}$$

For the cases (1), (2) and (4), clearly $\int_{[e_n, b_n + \epsilon_n)} (h_n^+ - u_2)^2 dm < \epsilon/M$. For case (3a), it follows from (2.13) that

$$\begin{aligned} \int_{[e_n, b_n + \epsilon_n)} (h_n^+ - u_2)^2 dm &= \int_{(b_n, b_n + \epsilon_n)} (h_n^+ - u_2)^2 dm \\ &\leq 2 \int_{(b_n, b_n + \epsilon_n)} (h_n^+ - h)^2 dm + 2 \int_{(b_n, b_n + \epsilon_n)} (h - u_2)^2 dm \\ &< \frac{2\epsilon}{M} + 2 \int_{(b_n, b_n + \epsilon_n)} (h - u_2)^2 dm. \end{aligned}$$

For the case (3b), we can deduce from (2.13) and (2.14) that

$$\int_{[e_n, b_n + \epsilon_n)} (h_n^+ - u_2)^2 dm < \frac{4\epsilon}{M} + 2 \int_{(b_n, b_n + \epsilon_n)} (h - u_2)^2 dm.$$

For the cases (3c) and (3d), from (2.14) and (2.15), we can obtain that

$$\int_{[e_n, b_n + \epsilon_n)} (h_n^+ - u_2)^2 dm < \frac{8\epsilon}{M} + 2 \int_{(b_n, b_n + \epsilon_n)} (h - u_2)^2 dm.$$

The term $\int_{(a_n - \epsilon'_n, e_n]} (h_n^- - u_2)^2 dm$ can be estimated similarly. Therefore, we can conclude that

$$\int_I |h_\epsilon - u_2|^2 dm < \frac{8\epsilon}{M} \cdot M + \frac{8\epsilon}{M} + 2 \int_I (h - u_2)^2 dm < 18\epsilon.$$

That completes the proof. \square

Finally, let h_ϵ be the function given in Lemma 2.9. We have

$$\mathcal{E}_1(h_\epsilon - u_2, h_\epsilon - u_2) = \mathcal{E}(h_\epsilon, h_\epsilon) + \|h_\epsilon - u_2\|_{L^2(I, m)}^2 < 24\epsilon.$$

and then, by setting $u_\epsilon := g_\epsilon + h_\epsilon \in \mathcal{F} \cap C_c(I)$, it holds that

$$\begin{aligned} \mathcal{E}_1(u_\epsilon - u, u_\epsilon - u) &\leq 2(\mathcal{E}_1(g_\epsilon - u_1, g_\epsilon - u_1) + \mathcal{E}_1(h_\epsilon - u_2, h_\epsilon - u_2)) \\ &< 60\epsilon. \end{aligned}$$

That completes the proof of sufficiency.

2.3. Effective intervals. Since the diffusion X in Theorem 2.1 has non-trivial motions only in each interval I_n ($n \geq 1$), we call the interval I_n with the scale function $\mathbf{s}_n \in \mathbf{S}_\infty(I_n)$ an *effective interval* of X or $(\mathcal{E}, \mathcal{F})$. All effective intervals of X (or $(\mathcal{E}, \mathcal{F})$) are $\{I_n : n \geq 1\}$ or $\{(I_n, \mathbf{s}_n) : n \geq 1\}$ if the scale functions are clarified. Clearly, X or $(\mathcal{E}, \mathcal{F})$ is determined by the effective intervals $\{(I_n, \mathbf{s}_n) : n \geq 1\}$ if the state space I and the symmetric measure m are given.

We give more remarks for the effective interval (I_n, \mathbf{s}_n) . Except for the case $I_n = \langle a_n, b_n \rangle$ shares the same open endpoint of I (i.e. (L0) or (R0)), the relation between I_n and \mathbf{s}_n is simple:

$$(2.16) \quad a_n \in I_n \Leftrightarrow \mathbf{s}_n(a_n) > -\infty, \quad b_n \in I_n \Leftrightarrow \mathbf{s}_n(b_n) < \infty.$$

This implies that a_n (resp. b_n) is a reflected endpoint if $a_n \in I_n$ (resp. $b_n \in I_n$) and an unapproachable endpoint if $a_n \notin I_n$ (resp. $b_n \notin I_n$) relative to X . For the case I_n has a common open endpoint of I , i.e. $I_n = \langle l, b_n \rangle$ and $I = \langle l, r \rangle$ (resp. $I_n = \langle a_n, r \rangle$ and $I = \langle l, r \rangle$), then $\mathbf{s}_n(l) > -\infty$ (resp. $\mathbf{s}_n(r) < \infty$) is allowed but l (resp. r) must be an absorbing endpoint relative to X (see (2.3)).

The example below gives us an intuitive understanding to the necessity of (2.16) for the regularity of $(\mathcal{E}, \mathcal{F})$.

Example 2.10. We first consider the absorbing Browning motion X on $I_1 = (0, 1)$. Precisely, it is associated with the regular Dirichlet form $(\frac{1}{2}\mathbf{D}, H_0^1(I_1))$ on $L^2(I_1)$, where

$$H_0^1(I_1) = \{u \in L^2(I_1) : u \text{ is absolutely continuous on } I_1, u' \in L^2(I_1), u(0) = u(1) = 0\},$$

$$\frac{1}{2}\mathbf{D}(u, v) = \frac{1}{2} \int_{I_1} u'(x)v'(x)dx, \quad u, v \in H_0^1(I_1).$$

Let $(B_t)_{t \geq 0}$ be the standard Brownian motion on \mathbb{R} and set $\zeta := \sigma_0 \wedge \sigma_1$ where σ_0 and σ_1 are the hitting times of $\{0\}$ and $\{1\}$ relative to $(B_t)_{t \geq 0}$. Clearly, ζ is the lifetime of X and

$$(2.17) \quad X_t = \begin{cases} B_t, & t < \zeta, \\ \partial, & t \geq \zeta, \end{cases}$$

where $(0, 1) \cup \partial$ is the one-point compactification of $(0, 1)$. As a rough observation, we may say ∂ coincides with $\{0, 1\}$. Thus $t \mapsto X_t$ is continuous on $[0, \infty)$ and X is a so-called Hunt process on $(0, 1)$.

Now let $I = [0, 1]$ and m be the Lebesgue measure on I . We treat the quadratic form $(\frac{1}{2}\mathbf{D}, H_0^1(I_1))$ on $L^2(I)$. Since $|\{0, 1\}| = 0$ and $u(0) = u(1) = 0$ for any $u \in H_0^1(I_1)$, it is really a Dirichlet form on $L^2(I)$ but not regular any more (notice that $H_0^1(I_1) \cap C(I)$ cannot separate 0 and 1). Nevertheless, Remark 2.11 after this example assures that $(\frac{1}{2}\mathbf{D}, H_0^1(I_1))$ is quasi-regular on $L^2(I)$ and its associated Markov process X^* is only a standard process on I . Let us explain why X^* is not a Hunt process. In fact, X^* still has an expression similar to (2.17):

$$X_t^* = \begin{cases} B_t, & t < \zeta, \\ \partial, & t \geq \zeta, \end{cases}$$

whereas ∂ is now an isolated point of I . We have $X_{\zeta-}^* = 0$ or 1 but $X_{\zeta}^* = \partial$ ($\neq 0$ or 1). Thus $t \mapsto X_t^*$ is only continuous on $[0, \zeta)$ and X^* is not a Hunt process on I .

Note that $I_1 = (0, 1)$ may be treated as an effective interval of $(\frac{1}{2}\mathbf{D}, H_0^1(I_1))$ on $L^2(I)$. The scale function on I_1 is the natural scale, i.e. $\mathbf{s}_1(x) = x$. Clearly, I_1 and \mathbf{s}_1 do not satisfy (2.16).

Remark 2.11. Fix a quasi-regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(E, m)$ (Cf. [14]). Assume E is a subset of a topological space E^* and extend m to a new measure m^* on E^* by $m^*(E^* \setminus E) = 0$. Then $(\mathcal{E}, \mathcal{F})$ is still a quasi-regular Dirichlet form on $L^2(E^*, m^*)$ and $E^* \setminus E$ is an \mathcal{E} -polar set relative to its associated Markov process.

The following example including the d -Bessel process for an integer $d \geq 2$ tells us a surprising fact: the presence of additional exceptional set might not break the regularity of the Dirichlet form. Notice that Example 2.10 is the contrary of this fact.

Example 2.12. Set $I_1 = (0, \infty)$ and consider an m -symmetric diffusion process Y on I_1 with a scale function \mathbf{s}_1 satisfying $\mathbf{s}_1(0) = -\infty$. For example, for some integer $d \geq 2$, let

$$(2.18) \quad \mathbf{s}_1(x) := \begin{cases} \log x, & d = 2, \\ \frac{x^{2-d}}{2-d}, & d \geq 3. \end{cases}$$

When $d \geq 3$, we have $\mathbf{s}_1(\infty) = 0$ and thus (R0) happens when $m([1, \infty)) < \infty$. The associated Dirichlet form of Y on $L^2(I_1, m)$ is given by

$$(2.19) \quad \mathcal{F} = \left\{ u \in L^2(I_1, m) : u \ll \mathbf{s}_1, \frac{du}{d\mathbf{s}_1} \in L^2(I_1, d\mathbf{s}_1), u(\infty) = 0 \text{ whenever (R0)} \right\},$$

$$\mathcal{E}(u, v) = \frac{1}{2} \int_{I_1} \frac{du}{d\mathbf{s}_1} \frac{dv}{d\mathbf{s}_1} d\mathbf{s}_1, \quad u, v \in \mathcal{F}.$$

Note that $(\mathcal{E}, \mathcal{F})$ is a regular Dirichlet form on $L^2(I_1, m)$ and Y is a regular diffusion on I_1 (Cf. [1, Proposition 2.2.10]). Particularly, when

$$(2.20) \quad m(dx) = x^{d-1}dx$$

and \mathbf{s}_1 is given by (2.18), Y is an equivalent version of the so-called d -Bessel process on $(0, \infty)$. It is well-known that the d -Bessel process (i.e. $(|B_t|)_{t \geq 0}$ for a d -dimensional Brownian motion $(B_t)_{t \geq 0}$) is usually defined as a regular diffusion on $[0, \infty)$, where 0 is the entrance boundary. So as in Example 2.10, we put $(\mathcal{E}, \mathcal{F})$ on the state space

$$I = [0, \infty).$$

Further assume that $m(0+) < \infty$ and then m can be naturally extended to a Radon measure, which we still denote it by m , on I (such as (2.20)). Clearly, $(\mathcal{E}, \mathcal{F})$ is still a Dirichlet form on $L^2(I, m)$. From Theorem 2.1, we can conclude that $(\mathcal{E}, \mathcal{F})$ is regular not only on $L^2(I_1, m)$ but also on $L^2(I, m)$. Particularly, the interval I_1 with a scale function \mathbf{s}_1 on it is an effective interval of regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(I, m)$ and $\{0\}$ is an \mathcal{E} -exceptional set. As an example, the associated Dirichlet form of d -Bessel process is regular on $L^2((0, \infty), m)$ and $L^2([0, \infty), m)$ at the same time.

Let us briefly explain why $(\mathcal{E}, \mathcal{F})$ is possibly regular on $L^2(I, m)$. Denote the associated diffusion process of $(\mathcal{E}, \mathcal{F})$ on $L^2(I, m)$ by Y^* . Then Y^* is a Hunt process on I . In fact, Y^* coincides with Y on I and $\{0\}$ is an \mathcal{E} -polar set relative to Y^* . The difference to X^* in Example 2.10 is that the boundary $\{0, 1\}$ is approachable for X or X^* but $\{0\}$ is not approachable for Y or Y^* since $\mathbf{s}_1(0) = -\infty$. In other words, Y is always staying in I before leaving from the boundary '0', no matter $\{0\}$ being or not being contained in the state space.

2.4. When $I = \mathbb{R}$. In this section, we shall discuss the special case of Theorem 2.1 for $I = \mathbb{R}$ which is of special importance. Let $J = \langle a, b \rangle \subset \mathbb{R}$ be an interval and $\mathbf{S}(J)$ be the class of scale functions on J as in §2.1. Since $I = \mathbb{R}$, the conditions **(A)** and **(B)** for a scale function $\mathbf{s} \in \mathbf{S}(J)$ could be written as

- (A_R)** When $a > -\infty$, $\mathbf{s}(a) = -\infty$ if and only if $a \notin J$;
- (B_R)** When $b < \infty$, $\mathbf{s}(b) = \infty$ if and only if $b \notin J$.

Then set

$$(2.21) \quad \mathbf{S}_\infty(J) := \{\mathbf{s} \in \mathbf{S}(J) : \mathbf{s} \text{ satisfies } \mathbf{(A_R)} \text{ and } \mathbf{(B_R)}\}.$$

Moreover, take $\mathbf{s} \in \mathbf{S}_\infty(J)$. The allowable situations (L0) and (R0) (roughly speaking, absorbing at $-\infty$ or ∞) are given as follows:

- (L_R)** $a = -\infty$, $\mathbf{s}(-\infty) > -\infty$ and $m((-\infty, 0]) < \infty$;
- (R_R)** $b = \infty$, $\mathbf{s}(\infty) < \infty$ and $m([0, \infty)) < \infty$.

Similarly, define a quadratic form $(\mathcal{E}^{(\mathbf{s})}, \mathcal{F}^{(\mathbf{s})})$ on $L^2(J, m_J)$ as follows:

$$(2.22) \quad \begin{aligned} \mathcal{F}^{(\mathbf{s})} &:= \left\{ u \in L^2(J, m_J) : u \ll \mathbf{s}, \frac{du}{d\mathbf{s}} \in L^2(J, d\mathbf{s}); \right. \\ &\quad \left. u(a) = 0 \text{ (resp. } u(b) = 0) \text{ whenever } \mathbf{(L_R)} \text{ (resp. } \mathbf{(R_R)})} \right\}, \end{aligned}$$

$$\mathcal{E}^{(\mathbf{s})}(u, v) := \frac{1}{2} \int_J \frac{du}{d\mathbf{s}} \frac{dv}{d\mathbf{s}} d\mathbf{s}, \quad u, v \in \mathcal{F}^{(\mathbf{s})}.$$

The following corollary is the special case of Theorem 2.1 for $I = \mathbb{R}$.

Corollary 2.13. *Let m be a fully supported Radon measure on \mathbb{R} . Then $(\mathcal{E}, \mathcal{F})$ is a regular and strongly local Dirichlet form on $L^2(\mathbb{R}, m)$ if and only if there exist a set of at most countable disjoint intervals $\{I_n = \langle a_n, b_n \rangle : I_n \subset \mathbb{R}, n \geq 1\}$ and a scale function $\mathbf{s}_n \in \mathbf{S}_\infty(I_n)$ for each $n \geq 1$ such that*

$$(2.23) \quad \begin{aligned} \mathcal{F} &= \left\{ u \in L^2(\mathbb{R}, m) : u|_{I_n} \in \mathcal{F}^{(\mathbf{s}_n)}, \sum_{n \geq 1} \mathcal{E}^{(\mathbf{s}_n)}(u|_{I_n}, u|_{I_n}) < \infty \right\}, \\ \mathcal{E}(u, v) &= \sum_{n \geq 1} \mathcal{E}^{(\mathbf{s}_n)}(u|_{I_n}, v|_{I_n}), \quad u, v \in \mathcal{F}, \end{aligned}$$

where for each $n \geq 1$, $(\mathcal{E}^{(\mathbf{s}_n)}, \mathcal{F}^{(\mathbf{s}_n)})$ is given by (2.22) with the scale function \mathbf{s}_n on I_n . Moreover, the intervals $\{I_n : n \geq 1\}$ and scale functions $\{\mathbf{s}_n : n \geq 1\}$ are uniquely determined, if the difference of order is ignored.

We end this section with a remark for the connection of Dirichlet forms in (2.4) and (2.23). Let $(\mathcal{E}, \mathcal{F})$ be the Dirichlet form (2.4) on $L^2(I, m)$, where $I = \langle l, r \rangle$ is an interval of \mathbb{R} , and X be its associated diffusion process on I . Further let $(\mathcal{E}^*, \mathcal{F}^*)$ be another regular and strongly local Dirichlet form on $L^2(I^*, m^*)$ associated with the diffusion X^* where I^* is another interval and m^* is a fully supported Radon measure on I^* . The effective intervals of $(\mathcal{E}^*, \mathcal{F}^*)$ are denoted by $\{(I_n^*, \mathbf{s}_n^*)\}$. We say $(\mathcal{E}, \mathcal{F})$ and $(\mathcal{E}^*, \mathcal{F}^*)$ are essentially equivalent if

$$\{(I_n, \mathbf{s}_n) : n \geq 1\} = \{(I_n^*, \mathbf{s}_n^*) : n \geq 1\}, \quad m|_{\cup_{n \geq 1} I_n} = m^*|_{\cup_{n \geq 1} I_n^*}.$$

This condition indicates the X^* coincides with X on each common effective interval (notice that outside the effective intervals, the diffusions have no motions). When $l (> -\infty)$ is a closed endpoint of I , i.e. $I = [l, r)$, we can extend m to a measure m^* on $(-\infty, r)$ by

$$m^*|_{(-\infty, l)} = dx|_{(-\infty, l)}, \quad m^*|_I := m|_I.$$

Then m^* is a fully supported Radon measure on $(-\infty, r)$ and $(\mathcal{E}, \mathcal{F})$ induces a new Dirichlet form on $L^2((-\infty, r), m^*)$:

$$(2.24) \quad \begin{aligned} \mathcal{F}^* &:= \{u \in L^2((-\infty, r), m^*) : u|_I \in \mathcal{F}\}, \\ \mathcal{E}^*(u, v) &:= \mathcal{E}(u|_I, v|_I), \quad u, v \in \mathcal{F}^*. \end{aligned}$$

Clearly, $(\mathcal{E}^*, \mathcal{F}^*)$ is regular by Theorem 2.1 and essentially equivalent to $(\mathcal{E}, \mathcal{F})$. When $I = (l, r)$, $m(l+) < \infty$ and (L0) does not appear, the same trick is also valid. Therefore, we can always find an essentially equivalent Dirichlet form of $(\mathcal{E}, \mathcal{F})$ on an extended open interval. Particularly, when m is a finite measure on I and (L0) and (R0) do not appear, this extended open interval can be taken as \mathbb{R} and $(\mathcal{E}, \mathcal{F})$ is essentially equivalent to a Dirichlet form $(\mathcal{E}^*, \mathcal{F}^*)$ on $L^2(\mathbb{R}, m^*)$, which can be characterized by Corollary 2.13. Roughly speaking, without loss of generality, we may always assume $I = \mathbb{R}$ when considering the questions about a regular and strongly local Dirichlet form associated with a symmetric linear diffusion.

3. WHEN \mathcal{F} CONTAINS SMOOTH FUNCTIONS

A typical way to construct a regular Dirichlet form is to start from a closable Markovian bilinear form on the domain C_c^∞ and then the closure of C_c^∞ with this form is a regular Dirichlet form, see [8, §3]. Let $(\mathcal{E}, \mathcal{F})$ be the regular Dirichlet form given in Theorem 2.1 and assume that

$$(3.1) \quad C_c^\infty(I) \subset \mathcal{F}.$$

What we are concerned of in this section is the representation of the closure of $C_c^\infty(I)$ in $(\mathcal{E}, \mathcal{F})$.

3.1. Regular Dirichlet subspaces. The theory of regular Dirichlet subspaces will be heavily used in this section. Given two regular Dirichlet forms $(\mathcal{E}^1, \mathcal{F}^1)$ and $(\mathcal{E}^2, \mathcal{F}^2)$ on $L^2(E, m)$, where E is a locally compact separable metric space and m is a fully supported Radon measure on E , $(\mathcal{E}^1, \mathcal{F}^1)$ is called a regular Dirichlet subspace of $(\mathcal{E}^2, \mathcal{F}^2)$ if

$$\mathcal{F}^1 \subset \mathcal{F}^2, \quad \mathcal{E}^2(u, v) = \mathcal{E}^1(u, v), \quad u, v \in \mathcal{F}^1.$$

Under the assumption (3.1), denote the closure of $C_c^\infty(I)$ in $(\mathcal{E}, \mathcal{F})$ by $\bar{\mathcal{F}}$ and define

$$\bar{\mathcal{E}}(u, v) := \mathcal{E}(u, v), \quad u, v \in \bar{\mathcal{F}}.$$

Then clearly $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ is a regular Dirichlet subspace of $(\mathcal{E}, \mathcal{F})$.

Basic properties about the regular Dirichlet subspaces are presented in §2 of [13]. Particularly, let $(\mathcal{E}, \mathcal{F})$ be a regular, irreducible and strongly local Dirichlet form on $L^2(J, m_J)$, where J is an open interval. Denote its scale function by \mathbf{s} . The following result is taken from [2] and [3] but we need a little more explanation for its proof.

Lemma 3.1. *Let $(\mathcal{E}, \mathcal{F}) = (\mathcal{E}^{(\mathbf{s})}, \mathcal{F}^{(\mathbf{s})})$ be given by (2.3), in which $I = J$. Then $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ is a regular Dirichlet subspace of $(\mathcal{E}, \mathcal{F})$ if and only if $(\bar{\mathcal{E}}, \bar{\mathcal{F}}) = (\mathcal{E}^{(\tilde{\mathbf{s}})}, \mathcal{F}^{(\tilde{\mathbf{s}})})$, where the scale function $\tilde{\mathbf{s}}$ satisfies*

$$(3.2) \quad d\tilde{\mathbf{s}} \ll d\mathbf{s}, \quad \frac{d\tilde{\mathbf{s}}}{d\mathbf{s}} = 0 \text{ or } 1, \quad d\mathbf{s}\text{-a.e.}$$

Proof. Before applying Theorem 4.1 of [3], we need to explain the following fact: if $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ is a regular Dirichlet subspace of $(\mathcal{E}, \mathcal{F})$, then $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ is strongly local and its associated diffusion process \tilde{X} is irreducible in the sense of (2.1). The strongly local property of $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ follows from Theorem 2.1 of [12]. Propositions 2.2 and 2.3 of [13] imply the irreducibility of \tilde{X} . The rest of the proof is clear. \square

3.2. Irreducible case. The irreducible case of $I = J$, where J is an open interval, is studied in [4]. Let $(\mathcal{E}, \mathcal{F})$ be the Dirichlet form on $L^2(J, m_J)$ in Lemma 3.1 and \mathbf{s} its scale function. The inverse function of \mathbf{s} is denoted by \mathbf{t} , i.e. $\mathbf{t} = \mathbf{s}^{-1}$. Then (3.1) is equivalent to the condition (Cf. [4, Lemma 2.1])

$$(3.3) \quad \mathbf{t} \text{ is absolutely continuous, and } \mathbf{t}' \in L^2_{\text{loc}}.$$

This assumption does not necessarily imply \mathbf{s} is absolutely continuous, but we may write the Lebesgue decomposition of $d\mathbf{s}$ as follows:

$$d\mathbf{s} = g(x)dx + \kappa, \quad \kappa \perp dx,$$

where $g \in L^1_{\text{loc}}(J)$. Note that $g > 0$ a.e. In fact, let $Z_g := \{x \in J \setminus H : g(x) = 0\}$, where H is the support of κ . Clearly, $|H| = 0$ and $d\mathbf{s}(Z_g) = \int_{Z_g} g(x)dx = 0$. Thus we must have $|Z_g| = 0$ since $dx \ll d\mathbf{s}$. From this fact, we can define the absolutely continuous part of \mathbf{s} as

$$\bar{\mathbf{s}}(x) := \int_e^x g(y)dy,$$

where e is a fixed point in J . Clearly, $\bar{\mathbf{s}}$ is a scale function on J , i.e. a strictly increasing and continuous function on J .

Assume (3.1) and denote the Dirichlet form induced by the closure of $C_c^\infty(J)$ by $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$. From Lemma 3.1, we know that $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ corresponds to a diffusion process with some scale function $\tilde{\mathbf{s}}$ satisfying (3.2). The following result is taken from [4, Theorem 2.2], which shows that $\tilde{\mathbf{s}} = \bar{\mathbf{s}}$.

Theorem 3.2. *Let $(\mathcal{E}, \mathcal{F})$ be in Lemma 3.1 and assume (3.1). Then the absolutely continuous part $\bar{\mathbf{s}}$ of \mathbf{s} is the scale function of regular Dirichlet form $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$. Particularly, $C_c^\infty(J)$ is a special standard core of $(\mathcal{E}, \mathcal{F})$ if and only if \mathbf{s} is absolutely continuous.*

3.3. General case. Now let us consider $I = \mathbb{R}$, m a fully supported Radon measure on \mathbb{R} and $(\mathcal{E}, \mathcal{F})$ a regular and strongly local Dirichlet form on $L^2(\mathbb{R}, m)$ whose effective intervals are $\{(I_n, \mathbf{s}_n) : n \geq 1\}$. For each n , e_n is a fixed point in the interior of I_n . The representation of $(\mathcal{E}, \mathcal{F})$ is given in Theorem 2.1. We shall point out that the assumption $I = \mathbb{R}$ is not essential, see Remark 3.9.

We first analyse the basic assumption (3.1). The following result follows from the fact that $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$ if and only if $f(x) = x \in \mathcal{F}_{\text{loc}}$.

Lemma 3.3. *Let $(\mathcal{E}, \mathcal{F})$ be a regular and strongly local Dirichlet form on $L^2(\mathbb{R}, m)$ with the effective intervals $\{(I_n, \mathbf{s}_n) : n \geq 1\}$. Then $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$ is equivalent to the following: $dx \ll d\mathbf{s}_n$ on I_n for any $n \geq 1$ and for any constant $L > 0$,*

$$(3.4) \quad \sum_{n \geq 1} \int_{I_n \cap (-L, L)} \left(\frac{dx}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n < \infty.$$

Proof. This equivalent condition (3.4) is clearly a concrete expression of $f \in \mathcal{F}_{\text{loc}}$ for $f(x) = x$. If $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$, then for any $L > 0$ we can take a function $f^L \in C_c^\infty(\mathbb{R})$ such that $f^L(x) = f(x)$ for any $x \in (-L, L)$. This implies $f \in \mathcal{F}_{\text{loc}}$. On the contrary, assume $f \in \mathcal{F}_{\text{loc}}$ and take $\varphi \in C_c^\infty(\mathbb{R})$ with $\text{supp}[\varphi] \subset (-L, L)$. Then it follows from (3.4) that $\varphi|_{I_n} \ll \mathbf{s}_n$ and

$$\sum_{n \geq 1} \mathcal{E}^{(\mathbf{s}_n)}(\varphi|_{I_n}, \varphi|_{I_n}) \leq \frac{\|\varphi'\|_\infty^2}{2} \sum_{n \geq 1} \int_{I_n \cap (-L, L)} \left(\frac{dx}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n < \infty.$$

Thus $\varphi \in \mathcal{F}$. That completes the proof. \square

Remark 3.4. We can deduce several corollaries from (3.4).

- (1) For each $n \geq 1$, the inverse $\mathbf{t}_n := \mathbf{s}_n^{-1}$ of \mathbf{s}_n is absolutely continuous.
- (2) If I_n is an internal interval (equivalently, a finite interval), then there exists a constant $L > 0$ such that $I_n \subset (-L, L)$. Thus

$$\int_{\mathbf{s}_n(I_n)} \mathbf{t}'_n(y)^2 dy < \infty.$$

- (3) If $I_n = \langle a_n, b_n \rangle$ is an infinite interval, i.e. $a_n = -\infty$ or $b_n = \infty$, then \mathbf{t}'_n is only a L^2_{loc} -function in its domain.

(4) If in addition each \mathbf{s}_n is absolutely continuous, then (3.4) is the same as

$$\sum_{n \geq 1} \int_{I_n \cap (-L, L)} \frac{1}{\mathbf{s}'_n(x)} dx < \infty.$$

We are now in a position to answer when $\bar{\mathcal{F}} = \mathcal{F}$. For this purpose, we need to give some definitions. Define two measures on \mathbb{R} :

$$(3.5) \quad \lambda_{\mathbf{s}} := \sum_{n \geq 1} d\mathbf{s}_n|_{I_n}, \quad \lambda_l := dx|_{(\bigcup_{n \geq 1} I_n)^c}.$$

In other words, $\lambda_{\mathbf{s}}$ is the sum of all measures associated to scale functions $\{\mathbf{s}_n : n \geq 1\}$ on effective intervals, not necessarily Radon, and λ_l is the restriction of Lebesgue measure to the rest part. Recall that e_n is a fixed point in the interior of I_n for $n \geq 1$.

Definition 3.5. For two effective intervals I_i and I_j , we say that I_i and I_j are *scale-connected*, or I_i is scale-connected to I_j , with respect to $\{\mathbf{s}_n : n \geq 1\}$, if

$$(3.6) \quad \lambda_{\mathbf{s}}([e_i, e_j]) < \infty, \quad \lambda_l([e_i, e_j]) = 0,$$

where $[e_i, e_j] := [e_j, e_i]$ if $e_j < e_i$. If I_i is not scale-connected to any other effective interval, we say that it is *scale-isolated*.

We write $I_i \stackrel{\mathbf{s}}{\sim} I_j$ to represent that I_i is scale-connected to I_j . Then one may easily check that $I_i \stackrel{\mathbf{s}}{\sim} I_i$, $I_i \stackrel{\mathbf{s}}{\sim} I_j$ implies $I_j \stackrel{\mathbf{s}}{\sim} I_i$, $I_i \stackrel{\mathbf{s}}{\sim} I_j$ and $I_j \stackrel{\mathbf{s}}{\sim} I_k$ imply $I_i \stackrel{\mathbf{s}}{\sim} I_k$. In other words, the scale-connection is an equivalence relation on all effective intervals $\{I_n : n \geq 1\}$.

Remark 3.6. Let I_i and I_j be two scale-connected effective intervals and $b_i < a_j$ (notice that $b_i = a_j$ is impossible). Intuitively, I_i is on the left side of I_j .

- (1) $\lambda_{\mathbf{s}}([e_i, e_j]) < \infty$ implies there are no open endpoints between I_i and I_j . More precisely, all the effective intervals between I_i and I_j are closed and b_i and a_j are also closed endpoint. However, a_i or b_j is not necessarily closed, see Example 3.8.
- (2) The first remark and $\lambda_l([e_i, e_j]) = 0$ indicate the effective intervals between I_i and I_j have a Cantor-type structure, see (for example) Example 3.8 (2, 3, 4).
- (3) If I_k is an effective interval between I_i and I_j , then I_k is scale-connected to I_i and I_j . Thus roughly speaking, the equivalence class of scale-connection separates all the effective intervals into several ‘continuous’ clusters.

Theorem 3.7. Let $(\mathcal{E}, \mathcal{F})$ be a strongly local and regular Dirichlet form on $L^2(\mathbb{R}, m)$ whose effective intervals are $\{(I_n, \mathbf{s}_n) : n \geq 1\}$ and assume $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$. Then $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}, \mathcal{F})$ if and only if the following conditions hold:

- (1) \mathbf{s}_n is absolutely continuous for any $n \geq 1$;
- (2) for any i , I_i is scale-isolated with respect to $\{\mathbf{s}_n : n \geq 1\}$.

Proof. We first prove the necessity. Assume that $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}, \mathcal{F})$.

For (1), fix n and recall that $I_n = \langle a_n, b_n \rangle$. The interior of I_n is denoted by $\overset{\circ}{I}_n = (a_n, b_n)$. Let us consider the part Dirichlet form $(\mathcal{E}_{\overset{\circ}{I}_n}^\circ, \mathcal{F}_{\overset{\circ}{I}_n}^\circ)$ of $(\mathcal{E}, \mathcal{F})$ on $\overset{\circ}{I}_n$. Clearly, $(\mathcal{E}_{\overset{\circ}{I}_n}^\circ, \mathcal{F}_{\overset{\circ}{I}_n}^\circ) = (\mathcal{E}_{\overset{\circ}{I}_n}^{(\mathbf{s}_n)}, \mathcal{F}_{\overset{\circ}{I}_n}^{(\mathbf{s}_n)})$, where the latter one is the part Dirichlet form of $(\mathcal{E}^{(\mathbf{s}_n)}, \mathcal{F}^{(\mathbf{s}_n)})$ on $\overset{\circ}{I}_n$. On the other hand, since $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}, \mathcal{F})$, it follows that $C_c^\infty(\overset{\circ}{I}_n)$ is a special standard core of $(\mathcal{E}_{\overset{\circ}{I}_n}^\circ, \mathcal{F}_{\overset{\circ}{I}_n}^\circ)$. Thus from Theorem 3.2 we can obtain that \mathbf{s}_n is absolutely continuous.

For (2), suppose that $I_1 = \langle a_1, b_1 \rangle$ and $I_2 = \langle a_2, b_2 \rangle$ are scale-connected. Without loss of generality, further assume $b_1 \leq a_2$. Let us consider the part Dirichlet form $(\mathcal{E}_J, \mathcal{F}_J)$ of $(\mathcal{E}, \mathcal{F})$ on $J = (a_1, b_2)$. Clearly, $C_c^\infty(J)$ is a special standard core of $(\mathcal{E}_J, \mathcal{F}_J)$ and $(\mathcal{E}_J, \mathcal{F}_J)$ is not irreducible. Take a fixed point $e \in J$. Since $I_1 \stackrel{\mathbf{s}}{\sim} I_2$, it follows that $\lambda_{\mathbf{s}}|_J$ is a Radon measure on J . Denote the induced increasing function of $\lambda_{\mathbf{s}}|_J$ by \mathbf{s}_J :

$$\mathbf{s}_J(x) := \int_e^x d\lambda_{\mathbf{s}}, \quad x \in J.$$

We assert that both \mathbf{s}_J and its inverse $\mathbf{t}_J := \mathbf{s}_J^{-1}$ are absolutely continuous, and $\mathbf{t}'_J \in L^2_{\text{loc}}(\mathbf{s}_J(J))$. In fact, the absolute continuity of \mathbf{s}_J follows from (1) and $d\mathbf{s}_J = \sum_{n \geq 1} d\mathbf{s}_n$ and the absolute

continuity of \mathbf{t}_J follows from Remark 3.4 (1) and $\lambda_l(J) = 0$. Take any closed interval $K \subset \mathbf{s}_J(J)$, it follows from Remark 3.4 (4) that

$$\int_K \mathbf{t}'_J(y)^2 dy = \int_{\mathbf{t}_J(K)} \mathbf{t}'_J \circ \mathbf{s}_J(x)^2 d\mathbf{s}_J(x) = \int_{\mathbf{t}_J(K)} \frac{1}{\mathbf{s}'_J(x)} dx < \infty.$$

From this assertion, (3.3) and Theorem 3.2, we can conclude that $C_c^\infty(J)$ is a special standard core of $(\mathcal{E}^{(\mathbf{s}_J)}, \mathcal{F}^{(\mathbf{s}_J)})$, where $(\mathcal{E}^{(\mathbf{s}_J)}, \mathcal{F}^{(\mathbf{s}_J)})$ is given by (2.3) with $I := J, \mathbf{s} := \mathbf{s}_J$. Now we shall verify that

$$\mathcal{E}^{(\mathbf{s}_J)}(\varphi, \varphi) = \mathcal{E}_J(\varphi, \varphi)$$

for any $\varphi \in C_c^\infty(J)$. Indeed, φ could be treated as a function in $C_c^\infty(\mathbb{R})$ with $\varphi|_{J^c} = 0$. Then we have

$$\begin{aligned} \mathcal{E}_J(\varphi, \varphi) &= \mathcal{E}(\varphi, \varphi) \\ &= \frac{1}{2} \sum_{n \geq 1} \int_{I_n \cap J} \left(\frac{d\varphi}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n(x) \\ &= \frac{1}{2} \sum_{n \geq 1} \int_{I_n \cap J} (\varphi'(x))^2 \frac{1}{\mathbf{s}'_n(x)} dx \\ &= \frac{1}{2} \int_J (\varphi'(x))^2 \frac{1}{\mathbf{s}'_J(x)} dx \\ &= \mathcal{E}^{(\mathbf{s}_J)}(\varphi, \varphi). \end{aligned}$$

This implies that $(\mathcal{E}^{(\mathbf{s}_J)}, \mathcal{F}^{(\mathbf{s}_J)}) = (\mathcal{E}_J, \mathcal{F}_J)$, which leads to a contradiction because $(\mathcal{E}^{(\mathbf{s}_J)}, \mathcal{F}^{(\mathbf{s}_J)})$ is irreducible but $(\mathcal{E}_J, \mathcal{F}_J)$ is not irreducible. Therefore, the necessity is proved.

Next, we turn to prove the sufficiency. Assume that (1) and (2) hold. Denote the closure of $C_c^\infty(\mathbb{R})$ in $(\mathcal{E}, \mathcal{F})$ by $\bar{\mathcal{F}}$ and set $\bar{\mathcal{E}}(u, v) := \mathcal{E}(u, v)$ for any $u, v \in \bar{\mathcal{F}}$. Then $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ is a regular and strongly local Dirichlet form on $L^2(\mathbb{R}, m)$. Applying Theorem 2.1 to $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$, denote the effective intervals of $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ by $\{(\bar{I}_n, \bar{\mathbf{s}}_n) : n \geq 1\}$. Write $\bar{I}_n = \langle \bar{a}_n, \bar{b}_n \rangle$. We need to show that

$$(3.7) \quad \{(I_n, \mathbf{s}_n) : n \geq 1\} = \{(\bar{I}_n, \bar{\mathbf{s}}_n) : n \geq 1\}.$$

We assert the following fact: the endpoints \bar{a}_n and \bar{b}_n are not the interior points of some I_k . Indeed, if $\bar{a}_n \in \overset{\circ}{I}_k$, consider the part Dirichlet forms of $(\mathcal{E}, \mathcal{F})$ and $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ on $J := \overset{\circ}{I}_k$. Denote the part Dirichlet forms by $(\mathcal{E}_J, \mathcal{F}_J)$ and $(\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J)$ respectively. Clearly, $(\mathcal{E}_J, \mathcal{F}_J)$ is irreducible but $(\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J)$ is not irreducible since $\bar{I}_n \cap J$ is a non-trivial invariant set of $(\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J)$. This is in contradiction with [13, Proposition 2.3 (3)] since $(\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J)$ is a regular Dirichlet subspace of $(\mathcal{E}_J, \mathcal{F}_J)$.

Given any \bar{I}_n , we shall prove that there exists some k such that $\bar{I}_n = I_k$ and $\bar{\mathbf{s}}_n = \mathbf{s}_k$. In other words,

$$(3.8) \quad \{(I_n, \mathbf{s}_n) : n \geq 1\} \supset \{(\bar{I}_n, \bar{\mathbf{s}}_n) : n \geq 1\}.$$

In fact, it follows from the above fact that for any j , either $\overset{\circ}{I}_j \cap \bar{I}_n = \emptyset$ or $\overset{\circ}{I}_j \subset \bar{I}_n$. Denote all the intervals with interior included in \bar{I}_n by

$$(3.9) \quad \{I_{k_j} : j \geq 1\}.$$

Note that they are disjoint with each other. We assert that any two I_{k_1} and I_{k_2} are scale-connected, thus (3.9) has at most one element by the condition (2). To this end, for any j , consider the part Dirichlet forms of $(\mathcal{E}, \mathcal{F})$ and $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ on $J := \overset{\circ}{I}_{k_j}$. We still denote them by $(\mathcal{E}_J, \mathcal{F}_J)$ and $(\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J)$. Clearly, we have

$$(\mathcal{E}_J, \mathcal{F}_J) = (\mathcal{E}_J^{(\mathbf{s}_{k_j})}, \mathcal{F}_J^{(\mathbf{s}_{k_j})}), \quad (\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J) = (\mathcal{E}_J^{(\bar{\mathbf{s}}_n)}, \mathcal{F}_J^{(\bar{\mathbf{s}}_n)})$$

on $L^2(J, m|_J)$. Since $C_c^\infty(J)$ is a special standard core of $(\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J)$, it follows from Theorem 3.2 that $\bar{\mathbf{s}}_n$ is absolutely continuous on J . Moreover, the inclusion $C_c^\infty(J) \subset \bar{\mathcal{F}}_J \subset \mathcal{F}_J$ indicates $\bar{\mathbf{s}}'_n > 0, \mathbf{s}'_{k_j} > 0$ a.e. on J by (3.3). Notice the \mathbf{s}_{k_j} is absolutely continuous by (1). On the other hand, $(\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J)$ is a regular Dirichlet subspace of $(\mathcal{E}_J, \mathcal{F}_J)$. This implies by Lemma 3.1

$$\frac{d\bar{\mathbf{s}}_n}{d\mathbf{s}_{k_j}} = \frac{\bar{\mathbf{s}}'_n}{\mathbf{s}'_{k_j}} = 0 \text{ or } 1, \quad \text{a.e. on } J.$$

Since $\bar{\mathbf{s}}'_n/\mathbf{s}'_{k_j} > 0$ a.e. on J , we can conclude that $d\bar{\mathbf{s}}_n = d\mathbf{s}_{k_j}$ on J . Particularly, $\lambda_{\bar{\mathbf{s}}} = d\bar{\mathbf{s}}_n$ on \bar{I}_n and thus

$$\lambda_{\bar{\mathbf{s}}}([e_{k_1}, e_{k_2}]) = \lambda_{\bar{\mathbf{s}}}([e_{k_1}, e_{k_2}]) < \infty.$$

Furthermore, for any $\varphi \in C_c^\infty((\bar{a}_n, \bar{b}_n))$, it follows from $\mathcal{E}(\varphi, \varphi) = \bar{\mathcal{E}}(\varphi, \varphi)$ that

$$\int_{\bar{a}_n}^{\bar{b}_n} \left(\frac{d\varphi}{d\bar{\mathbf{s}}_n} \right)^2 d\bar{\mathbf{s}}_n = \sum_j \int_{I_{k_j}} \left(\frac{d\varphi}{d\mathbf{s}_{k_j}} \right)^2 d\mathbf{s}_{k_j} = \sum_j \int_{I_{k_j}} \left(\frac{d\varphi}{d\bar{\mathbf{s}}_n} \right)^2 d\bar{\mathbf{s}}_n.$$

Therefore, we have

$$\lambda_l(\bar{I}_n) = |\bar{I}_n \setminus (\bigcup_j I_{k_j})| = 0$$

and it follow that I_{k_1} is scale-connected to I_{k_2} . In addition, the fact that $\lambda_l(\bar{I}_n) = 0$ also implies (3.9) has at least one element. Hence combining together, (3.9) has exactly one element, which is denoted by $\overset{\circ}{I}_k$, and $\overset{\circ}{I}_k = (\bar{a}_n, \bar{b}_n)$ and $d\mathbf{s}_k = d\bar{\mathbf{s}}_n$. Notice that \mathbf{s}_k and $\bar{\mathbf{s}}_n$ satisfy **(A)** and **(B)** (see (2.2)), and the fixed point e_k in $\overset{\circ}{I}_k = (\bar{a}_n, \bar{b}_n)$ for \mathbf{s}_k and $\bar{\mathbf{s}}_n$ (i.e. $\mathbf{s}_k(e_k) = \bar{\mathbf{s}}_n(e_k) = 0$) could be taken as the same one. We then conclude that $I_k = \bar{I}_n$ and $\mathbf{s}_k = \bar{\mathbf{s}}_n$.

Finally, suppose that some (I_k, \mathbf{s}_k) is not contained in the right side of (3.7). Let us consider the the part Dirichlet forms $(\mathcal{E}_J, \mathcal{F}_J)$ and $(\bar{\mathcal{E}}_J, \bar{\mathcal{F}}_J)$ of $(\mathcal{E}, \mathcal{F})$ and $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ on $J := \overset{\circ}{I}_k$. Note that $C_c^\infty(J) \subset \bar{\mathcal{F}}_J \subset \mathcal{F}_J$ and for any $\varphi \in C_c^\infty(J)$, $\mathcal{E}_J(\varphi, \varphi) = \bar{\mathcal{E}}(\varphi, \varphi)$. However, $\bar{\mathcal{E}}_J(\varphi, \varphi) = 0$ but

$$\mathcal{E}_J(\varphi, \varphi) = \frac{1}{2} \int_J \left(\frac{d\varphi}{d\mathbf{s}_k} \right)^2 d\mathbf{s}_k,$$

which leads to a contradiction. That completes the proof. \square

Example 3.8. As characterized in Theorem 2.1, the Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(\mathbb{R}, m)$ is constituted by a sequence (possibly, none, finite, or infinite) of effective intervals with scale functions $\{(I_n, \mathbf{s}_n) : n \geq 1\}$.

- (1) The first example is very simple. Let

$$\begin{aligned} I_1 &= (-\infty, 0], \quad \mathbf{s}_1(x) = x; \\ I_2 &= (0, \infty), \quad \mathbf{s}_2(x) = -\exp\{1/x\}. \end{aligned}$$

Then $\{(I_i, \mathbf{s}_i) : i = 1, 2\}$ satisfies all the conditions including (1) and (2) in Theorem 3.7. Particularly, I_1 and I_2 are scale-isolated with respect to $\{\mathbf{s}_1, \mathbf{s}_2\}$. As a result, $C_c^\infty(\mathbb{R})$ is a special standard core of the associated Dirichlet form.

- (2) The second example is taken from [13, Example 3.20]. Let K be the standard Cantor set in $[0, 1]$. Set $U := K^c$ and write U as a union of disjoint open intervals:

$$U = \bigcup_{n \geq 1} (a_n, b_n),$$

where $(a_1, b_1) = (-\infty, 0)$, $(a_2, b_2) = (1, \infty)$. Let $I_1 := (-\infty, 0]$, $I_2 := [1, \infty)$, $I_n := [a_n, b_n]$ for any $n \geq 3$. For each n , let $\mathbf{s}_n(x) = x$ on I_n . Then the associated Dirichlet form $(\mathcal{E}, \mathcal{F})$ is a so-called regular Dirichlet extension of one-dimensional Brownian motion (Cf. [13]) and the associated diffusion process X is a reflected Brownian motion on each interval I_n . One may easily check that I_1 is scale-connected to I_2 and thus for any i, j , I_i is scale-connected to I_j . Hence $C_c^\infty(\mathbb{R})$ is not a special standard core of $(\mathcal{E}, \mathcal{F})$ by Theorem 3.7. Indeed, as indicated in [13], the closure of $C_c^\infty(\mathbb{R})$ in $(\mathcal{E}, \mathcal{F})$ corresponds to the one-dimensional Brownian motion.

- (3) We consider the intervals $\{I_n : n \geq 1\}$ in the second example but redefine the scale functions as follows: $\mathbf{s}_1(x) = x$, $\mathbf{s}_2(x) = x$ and

$$\mathbf{s}_n(x) := \frac{x}{|b_n - a_n|}, \quad x \in I_n, n \geq 3.$$

Then we have $d\mathbf{s}_n(I_n) = 1$ for any $n \geq 3$. Clearly, $(\mathcal{E}, \mathcal{F})$ satisfies (3.4) and thus $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$. Notice that for any $i \neq j$, there are infinite effective intervals between I_i and I_j . This implies that each I_i is scale-isolated since $\lambda_{\mathbf{s}}([e_i, e_j]) = \infty$. Therefore from Theorem 3.7 it follows that $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}, \mathcal{F})$.

- (4) Finally, we consider an example similar to the second one but replace K by a generalized Cantor set. In other words, K has a Cantor-type structure but the Lebesgue measure of K is positive. We refer to [5, §1.5] for the details to construct the generalized Cantor set. The effective intervals with the scale functions are given by the same method as the second example. Then (3.4) holds and $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$. For any $i \neq j$, clearly $\lambda_s([e_i, e_j]) < \infty$ since $d\mathbf{s}_n$ is nothing but the Lebesgue measure. However, one may easily show that $\lambda_l([e_i, e_j]) > 0$ from the structure of generalized Cantor set. This indicates that each I_i is scale-isolated and therefore, $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}, \mathcal{F})$.

We end this section with a remark to explain that the setting $I = \mathbb{R}$ is not essential as we pointed out in the beginning of this section.

Remark 3.9. When I is an open interval, all the considerations above are robust and only very few details need to be modified. For example, $(-L, L)$ in (3.4) could be replaced by a compact sub-interval of I .

If I is not open, we can apply the trick in §2.3 to find an essentially equivalent Dirichlet form $(\mathcal{E}^*, \mathcal{F}^*)$ on a larger open interval and transfer the questions to the new Dirichlet form. For simplicity, let us consider the following example: $I = [0, \infty)$ and $(\mathcal{E}, \mathcal{F})$ is a regular and strongly local Dirichlet form on $L^2(I, m)$. Let $(\mathcal{E}^*, \mathcal{F}^*)$ on $L^2(\mathbb{R}, m^*)$ be given by (2.24), which is an essentially equivalent Dirichlet form of $(\mathcal{E}, \mathcal{F})$. Clearly, $C_c^\infty(\mathbb{R}) \subset \mathcal{F}^*$ is equivalent to

$$C_c^\infty(I) = \{\varphi|_I : \varphi \in C_c^\infty(\mathbb{R})\} \subset \mathcal{F}.$$

We assert that $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}^*, \mathcal{F}^*)$ if and only if $C_c^\infty(I)$ is a special standard core of $(\mathcal{E}, \mathcal{F})$. The necessity is clear. For the sufficiency, assume that $C_c^\infty(I)$ is a special standard core of $(\mathcal{E}, \mathcal{F})$, we need to prove $C_c^\infty(\mathbb{R})$ is a core of $(\mathcal{E}^*, \mathcal{F}^*)$. Consider the part Dirichlet form $(\mathcal{E}_0, \mathcal{F}_0)$ of $(\mathcal{E}, \mathcal{F})$ on $I = (0, \infty)$. Note that $C_c^\infty((0, \infty))$ is a special standard core of $(\mathcal{E}_0, \mathcal{F}_0)$. Mimicking Theorem 3.7, we can deduce that the effective intervals of $(\mathcal{E}_0, \mathcal{F}_0)$ satisfy the conditions (1) and (2) of Theorem 3.7. Notice that the effective intervals are almost the same as those of $(\mathcal{E}, \mathcal{F})$ (except for the case $[0, b)$ is an effective interval of $(\mathcal{E}, \mathcal{F})$ but the corresponding one of $(\mathcal{E}_0, \mathcal{F}_0)$ is $(0, b)$). One easily sees that $(\mathcal{E}, \mathcal{F})$ (hence $(\mathcal{E}^*, \mathcal{F}^*)$) also satisfies the conditions (1) and (2) of Theorem 3.7. Applying Theorem 3.7 to $(\mathcal{E}^*, \mathcal{F}^*)$ again, we conclude that $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}^*, \mathcal{F}^*)$.

3.4. Recovering Hamza's theorem. A class of strongly local and regular Dirichlet forms on $L^2(\mathbb{R}, m)$ was introduced by Hamza [9], which starts from an energy form defined on $C_c^\infty(\mathbb{R})$. For convenience, we introduce a stronger version which is stated in [8]. Let ν be a positive Radon measure on \mathbb{R} and define the following quadratic form:

$$(3.10) \quad \begin{aligned} \mathcal{D}(\mathcal{E}) &:= C_c^\infty(\mathbb{R}), \\ \mathcal{E}(u, v) &:= \frac{1}{2} \int_{\mathbb{R}} u'(x)v'(x)\nu(dx), \quad u, v \in \mathcal{D}(\mathcal{E}). \end{aligned}$$

The aim is to find some conditions on ν , under which $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is closable on $L^2(\mathbb{R}, m)$. Given a non-negative function $a(x)$ on \mathbb{R} , we say $x \in \mathbb{R}$ is a regular point of a if for some $\epsilon > 0$,

$$(3.11) \quad \int_{x-\epsilon}^{x+\epsilon} \frac{1}{a(\xi)} d\xi < \infty.$$

The set of all regular points of a are denoted by $R(a)$ and $S(a) := \mathbb{R} \setminus R(a)$ is called the singular set of a . The following theorem is taken from [8, Theorem 3.1.6] and we give an alternative proof of the necessity, and the sufficiency is left for further discussions afterwards.

Theorem 3.10. *The form (3.10) is closable on $L^2(\mathbb{R}, m)$ if and only if the following conditions are satisfied:*

- (1) ν is absolutely continuous,
- (2) the density function a ($\nu(dx) = a(x)dx$) vanished a.e. on its singular set $S(a)$.

Proof of necessity. Denote the closure of $\mathcal{D}(\mathcal{E})$ relative to \mathcal{E}_1 by \mathcal{F} . Then $(\mathcal{E}, \mathcal{F})$ is a regular and strongly local Dirichlet form on $L^2(\mathbb{R}, m)$ and $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}, \mathcal{F})$. Applying Theorem 2.1 to $(\mathcal{E}, \mathcal{F})$, denote its effective intervals by $\{(I_n, \mathbf{s}_n) : n \geq 1\}$. It follows

from Lemma 3.3 and Theorem 3.7 that \mathbf{s}_n is absolutely continuous and $\mathbf{s}'_n > 0$ a.e. on I_n . Moreover, for any $\varphi \in C_c^\infty(\mathbb{R})$, we have

$$(3.12) \quad \mathcal{E}(\varphi, \varphi) = \frac{1}{2} \sum_{n \geq 1} \int_{I_n} \left(\frac{d\varphi}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n = \frac{1}{2} \sum_{n \geq 1} \int_{I_n} (\varphi'(x))^2 \frac{1}{\mathbf{s}'_n(x)} dx.$$

Define a measure μ as follows:

$$\mu|_{\{\mathbb{R} \setminus \bigcup_{n \geq 1} I_n\}} = 0, \quad \mu|_{I_n} := \frac{1}{\mathbf{s}'_n(x)} dx, \quad n \geq 1.$$

Then Lemma 3.3 and Remark 3.4 (4) imply that μ is a Radon measure on \mathbb{R} and from (3.10) and (3.12) we can deduce that $\mu = \nu$. Particularly, ν is absolutely continuous, i.e. (1) holds. From the fact each I_i is scale-isolated (see Theorem 3.7), we conclude that the regular set of the density function a of ν is

$$R(a) = \bigcup_{n \geq 1} \overset{\circ}{I}_n,$$

where $\overset{\circ}{I}_n$ is the interior of I_n . Clearly, the density function a vanishes a.e. on $S(a)$. That completes the proof. \square

Assume the quadratic form (3.10) satisfies the two conditions in Theorem 3.10 and denote the closure of $\mathcal{D}(\mathcal{E})$ relative to \mathcal{E}_1 by \mathcal{F} . Clearly, $(\mathcal{E}, \mathcal{F})$ is a strongly local and regular Dirichlet form on $L^2(\mathbb{R})$, and $C_c^\infty(\mathbb{R})$ is a special standard core of $(\mathcal{E}, \mathcal{F})$. The main purpose of rest part is to identify the effective intervals of $(\mathcal{E}, \mathcal{F})$, so that we can obtain an explicit structure of the associated diffusion process.

We first note that the condition that ν is Radon plays a similar role to (3.1) since it guarantees that the integration in (3.10) is finite. On the other hand, from (3.11) we can easily deduce that the regular set $R(a)$ is open. Thus we can write $R(a)$ as a union of disjoint open intervals:

$$(3.13) \quad R(a) = \bigcup_{n \geq 1} (a_n, b_n).$$

For each n take a fixed point $e_n \in (a_n, b_n)$ and define

$$(3.14) \quad \mathbf{s}_n(x) := \int_{e_n}^x \frac{1}{a(\xi)} d\xi, \quad x \in (a_n, b_n).$$

We assert \mathbf{s}_n is an absolutely continuous scale function on (a_n, b_n) , i.e. \mathbf{s}_n is absolutely continuous and strictly increasing. In fact, one may easily verify that $|\mathbf{s}_n(x)| < \infty$ by (3.11) and Heine-Borel theorem. It follows that $1/a \in L^1_{\text{loc}}((a_n, b_n))$ and \mathbf{s}_n is absolutely continuous. Clearly, \mathbf{s}_n is increasing, and it is actually strictly increasing. In fact suppose that $\mathbf{s}_n(x) = \mathbf{s}_n(y)$ for $a_n < x < y < b_n$. Then $a(\xi) = \infty$ for any $\xi \in (x, y)$, which contradicts that $\nu(d\xi) = a(\xi)d\xi$ is Radon.

Naturally, we can extend the definition of \mathbf{s}_n to

$$\mathbf{s}_n(a_n) := \lim_{x \downarrow a_n} \mathbf{s}_n(x) (\geq -\infty), \quad \mathbf{s}_n(b_n) := \lim_{x \uparrow b_n} \mathbf{s}_n(x) (\leq \infty).$$

Set

$$(3.15) \quad I_n := \langle a_n, b_n \rangle,$$

where $a_n \in I_n$ (resp. $b_n \in I_n$) if and only if $a_n > -\infty, \mathbf{s}_n(a_n) > -\infty$ (resp. $b_n < \infty, \mathbf{s}_n(b_n) < \infty$). Then $\{I_n : n \geq 1\}$ is mutually disjoint. Indeed, suppose that $I_i = \langle a_i, b_i \rangle, I_j = [a_j, b_j]$ and $b_i = a_j$. Then $\mathbf{s}_i(b_i) < \infty$ and $\mathbf{s}_j(a_j) > -\infty$ imply the common endpoint $b_i = a_j$ is a regular point of a . However, (3.13) tells us $b_i = a_j \notin R(a)$, which leads to a contradiction.

Corollary 3.11. *Let $(\mathcal{E}, \mathcal{F})$ be the closure of (3.10) on $L^2(\mathbb{R}, m)$ under the conditions of Theorem 3.10. The interval I_n and the scale function \mathbf{s}_n on I_n are defined by (3.15) and (3.14) respectively for each $n \geq 1$. Then $(\mathcal{E}, \mathcal{F})$ can be expressed as (2.4) with the effective intervals $\{(I_n, \mathbf{s}_n) : n \geq 1\}$.*

Proof. Let $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ be the Dirichlet form given by (2.4) with effective intervals $\{(I_n, \mathbf{s}_n) : n \geq 1\}$. We need to show $(\mathcal{E}, \mathcal{F}) = (\bar{\mathcal{E}}, \bar{\mathcal{F}})$.

We first assert that $C_c^\infty(\mathbb{R}) \subset \bar{\mathcal{F}}$. In fact, since $\mathbf{s}'_n = 1/a > 0$ a.e. on I_n , it follows that $\mathbf{t}_n := \mathbf{s}_n^{-1}$ is absolutely continuous. By Lemma 3.3, Remark 3.4 (4) and

$$\sum_{n \geq 1} \int_{I_n \cap (-L, L)} \frac{1}{\mathbf{s}'_n(x)} dx = \sum_{n \geq 1} \int_{I_n \cap (-L, L)} a(x) dx = \nu((-L, L)) < \infty, \quad \forall L > 0,$$

it can be seen that $C_c^\infty(\mathbb{R}) \subset \bar{\mathcal{F}}$. Moreover, for any $\varphi \in C_c^\infty(\mathbb{R})$, from Remark 3.4 (4) and the second condition of Theorem 3.10, we can deduce that

$$\begin{aligned} \bar{\mathcal{E}}(\varphi, \varphi) &= \frac{1}{2} \sum_{n \geq 1} \int_{I_n} \left(\frac{d\varphi}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n \\ &= \frac{1}{2} \sum_{n \geq 1} \int_{I_n} \varphi'(x)^2 \frac{1}{\mathbf{s}'_n(x)} dx \\ &= \frac{1}{2} \sum_{n \geq 1} \int_{I_n} \varphi'(x)^2 a(x) dx \\ &= \frac{1}{2} \int_{\mathbb{R}} \varphi'(x)^2 \nu(dx) \\ &= \mathcal{E}(\varphi, \varphi). \end{aligned}$$

Thus we need only to show that $C_c^\infty(\mathbb{R})$ is a special standard core of $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$. By Theorem 3.7, it suffices to prove that each I_i is scale-isolated with respect to $\{\mathbf{s}_n : n \geq 1\}$. Suppose that I_i is scale-connected to I_j for some $j \neq i$. On one hand, there exists some point x between I_i and I_j , which is not a regular point of a by (3.13). On the other hand, the measure $\lambda_{\mathbf{s}}$ defined by (3.5)

$$\lambda_{\mathbf{s}} = (1/a(\xi)) \cdot d\xi|_{\cup_{n \geq 1} I_n}.$$

Thus from $\lambda_l([e_i, e_j]) = 0$, we can obtain that for some $\epsilon > 0$ such that $(x - \epsilon, x + \epsilon) \subset [e_i, e_j]$,

$$\int_{x-\epsilon}^{x+\epsilon} \frac{1}{a(\xi)} d\xi = \int_{x-\epsilon}^{x+\epsilon} d\lambda_{\mathbf{s}} \leq \lambda_{\mathbf{s}}([e_i, e_j]) < \infty.$$

In other words, $x \in R(a)$, which leads to a contradiction. That completes the proof. \square

3.5. Closure of C_c^∞ . Let $(\mathcal{E}, \mathcal{F})$ be a strongly local and regular Dirichlet form on $L^2(\mathbb{R}, m)$ with the effective intervals $\{(I_n, \mathbf{s}_n) : n \geq 1\}$ and assume that

$$C_c^\infty(\mathbb{R}) \subset \mathcal{F}.$$

We still denote the Dirichlet form associated with the closure of $C_c^\infty(\mathbb{R})$ in $(\mathcal{E}, \mathcal{F})$ by $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$. This section is devoted to explore the structure of $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$, which is not necessarily accordant with $(\mathcal{E}, \mathcal{F})$. Since $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ is also a regular and strongly local Dirichlet form on $L^2(\mathbb{R}, m)$, it follows from Theorem 2.1 that $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ admits the expression of (2.4) with another sequence of effective intervals with scale functions, which is our essential target.

Before presenting the main theorem, we need to do some preparation. As the irreducible case in §3.2, the scale function \mathbf{s}_n is not necessarily absolutely continuous. However we may write its Lebesgue decomposition as follows:

$$d\mathbf{s}_n = g_n(x)dx + \kappa_n, \quad \kappa_n \perp dx.$$

Mimicking §3.2, we can define the absolutely continuous part of \mathbf{s}_n by

$$\bar{\mathbf{s}}_n(x) := \int_{e_n}^x g_n(y)dy, \quad x \in I_n.$$

We know that $\bar{\mathbf{s}}_n$ is an absolutely continuous scale function on I_n . One may guess the effective intervals of $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ are $\{(I_n, \bar{\mathbf{s}}_n) : n \geq 1\}$. But this is not correct. A counterexample is Example 3.8 (2), in which \mathbf{s}_n is absolutely continuous but $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ corresponds to the one-dimensional Brownian motion. Another example below shows that $\bar{\mathbf{s}}_n$ and I_n may not match, i.e. $\bar{\mathbf{s}}_n \notin \mathbf{S}_\infty(I_n)$, where $\mathbf{S}_\infty(I_n)$ is given by (2.2).

Example 3.12. Let $I := (0, 1]$. Further let \mathbf{c} be the standard Cantor function on $[0, 1]$, particularly $\mathbf{c}(0) = 0, \mathbf{c}(1) = 1$. Then \mathbf{c} induces a singular (i.e. $d\mathbf{c} \perp dx$) probability measure $d\mathbf{c}$ on $(0, 1)$. For each integer $p \geq 1$, consider the map

$$f_p : (0, 1) \rightarrow \left(\frac{1}{2^p}, \frac{1}{2^{p-1}} \right), \quad x \mapsto \frac{1+x}{2^p}$$

and denote the image measure of $d\mathbf{c}$ relative to f_p by $d\mathbf{c}_p := d\mathbf{c} \circ f_p^{-1}$. Then $d\mathbf{c}_p$ is a probability measure on $(\frac{1}{2^p}, \frac{1}{2^{p-1}})$. Clearly, $\sum_{p \geq 1} d\mathbf{c}_p + dx$ defines a Radon measure on $I = (0, 1]$. Define \mathbf{s} to be the increasing function associated to this measure, in other words, $\mathbf{s}(1) = 1$ and

$$(3.16) \quad d\mathbf{s} = dx + \sum_{p \geq 1} d\mathbf{c}_p.$$

Note that $\mathbf{s}(0) := \lim_{x \downarrow 0} \mathbf{s}(x) = -\infty$. Thus $\mathbf{s} \in \mathbf{S}_\infty(I)$. On the other hand, since $d\mathbf{c}_p \perp dx$, it follows that (3.16) is the Lebesgue decomposition of $d\mathbf{s}$. This implies that $d\bar{\mathbf{s}}(x) = dx$ and $\bar{\mathbf{s}} \notin \mathbf{S}_\infty(I)$.

From these examples, we can see that both intervals and scale functions will change. Recall the definition of scale-connection for the effective intervals in Definition 3.5. Similarly we use the new family of scale functions $\{\bar{\mathbf{s}}_n\}$ to define scale-connection. Precisely set

$$\lambda_{\bar{\mathbf{s}}} := \sum_{n \geq 1} d\bar{\mathbf{s}}_n|_{I_n}.$$

Two effective intervals I_i and I_j are scale-connected with respect to $\{\bar{\mathbf{s}}_n : n \geq 1\}$ if

$$\lambda_{\bar{\mathbf{s}}}([e_i, e_j]) < \infty, \quad \lambda_l([e_i, e_j]) = 0.$$

Note that $\lambda_{\bar{\mathbf{s}}} \leq \lambda_{\mathbf{s}}$. It follows that the scale-connection with respect to $\{\mathbf{s}_n : n \geq 1\}$ implies the scale-connection with respect to $\{\bar{\mathbf{s}}_n : n \geq 1\}$ but not vice versa. We write $I_i \stackrel{\bar{\mathbf{s}}}{\sim} I_j$ to represent that I_i is scale-connected to I_j with respect to $\{\bar{\mathbf{s}}_n : n \geq 1\}$. Clearly, ' $\stackrel{\bar{\mathbf{s}}}{\sim}$ ' is also an equivalence relation on all effective intervals $\{I_n : n \geq 1\}$. Denote all the equivalence classes induced by $\stackrel{\bar{\mathbf{s}}}{\sim}$ by

$$\bar{\mathcal{I}}_k = \{I_{k_j} : j \geq 1\}, \quad k \geq 1,$$

where k is the index of the equivalence class and $\{k_j : j \geq 1\}$ are the indexes of the effective intervals in k -th equivalence class. Note that both numbers of equivalence classes and effective intervals in each equivalence class may be finite or infinite. As described in Remark 3.6 (3), $\bar{\mathcal{I}}_k$ also looks like a 'continuous' cluster.

Fix k , and define

$$\bar{\mathbf{a}}_k := \inf\{x \in I_{k_j} : j \geq 1\}, \quad \bar{\mathbf{b}}_k := \sup\{x \in I_{k_j} : j \geq 1\}.$$

In other words, $[\bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k]$ is the smallest closed interval, which contains all the intervals in the equivalence class $\bar{\mathcal{I}}_k$ (the notation $[\bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k]$ is a little abused since $\bar{\mathbf{a}}_k$ or $\bar{\mathbf{b}}_k$ may be infinity). One may easily check that $\lambda_{\bar{\mathbf{s}}}|_{(\bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k)}$ is a Radon measure on $(\bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k)$. Then we can take a fixed point $\bar{\mathbf{e}}_k \in (\bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k)$ and define

$$\bar{\mathbf{s}}_k(x) := \int_{\bar{\mathbf{e}}_k}^x d\lambda_{\bar{\mathbf{s}}}, \quad x \in [\bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k].$$

Clearly, $\bar{\mathbf{s}}_k$ is a scale function on $[\bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k]$ and $\bar{\mathbf{s}}_k(\bar{\mathbf{a}}_k) \geq -\infty, \bar{\mathbf{s}}_k(\bar{\mathbf{b}}_k) \leq \infty$. Further define

$$\bar{\mathcal{I}}_k := \langle \bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k \rangle,$$

where $\bar{\mathbf{a}}_k \in \bar{\mathcal{I}}_k$ (resp. $\bar{\mathbf{b}}_k \in \bar{\mathcal{I}}_k$) if and only if $\bar{\mathbf{a}}_k > -\infty, \bar{\mathbf{s}}_k(\bar{\mathbf{a}}_k) > -\infty$ (resp. $\bar{\mathbf{b}}_k < \infty, \bar{\mathbf{s}}_k(\bar{\mathbf{b}}_k) < \infty$). We need to check that these intervals $\{\bar{\mathcal{I}}_k : k \geq 1\}$ are mutually disjoint. Indeed, if $\bar{\mathbf{b}}_1 = \bar{\mathbf{a}}_2 \in \bar{\mathcal{I}}_1 \cap \bar{\mathcal{I}}_2$, then any interval in $\bar{\mathcal{I}}_1$ is scale-connected to any interval in $\bar{\mathcal{I}}_2$ with respect to $\{\bar{\mathbf{s}}_n : n \geq 1\}$. In other words, $\bar{\mathcal{I}}_1 = \bar{\mathcal{I}}_2$, which contradicts the definition of equivalence classes.

Intuitively, as described above we merge all the intervals in $\bar{\mathcal{I}}_k$ into a new one $\bar{\mathcal{I}}_k$ (notice that $\lambda_l(\bar{\mathcal{I}}_k) = 0$) and obtain a new scale function $\bar{\mathbf{s}}_k$ by the sum of $\{\bar{\mathbf{s}}_{k_j} : j \geq 1\}$. Now we give the following definition.

Definition 3.13. The pair $(\bar{\mathcal{I}}_k, \bar{\mathbf{s}}_k)$ is called the C^∞ -merging of the equivalence class $\bar{\mathcal{I}}_k$ for any $k \geq 1$.

In Example 3.8 (2), $\bar{\mathbf{s}}_n = \mathbf{s}_n$ and all the effective intervals are scale-connected to each other. Thus there is only one equivalence class $\tilde{\mathcal{I}}_1$ and we have $\bar{\mathbf{a}}_1 = -\infty, \bar{\mathbf{b}}_1 = \infty$. Take the fixed point $\bar{\mathbf{e}}_1 = 0$ and we can obtain $\bar{\mathbf{s}}_1(x) = x$ and $\bar{\mathcal{I}}_1 = \mathbb{R}$. Recall the $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}}) = (\frac{1}{2}\mathbf{D}, H^1(\mathbb{R}))$, which is the associated Dirichlet form of one-dimensional Brownian motion. Thus the C^∞ -merging $(\bar{\mathcal{I}}_1, \bar{\mathbf{s}}_1)$ of $\tilde{\mathcal{I}}_1$ is exactly the effective interval of $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$. The following theorem tells us for general cases all C^∞ -merging constitute actually the effective intervals of $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$.

Theorem 3.14. *Let $(\mathcal{E}, \mathcal{F})$ be a strongly local and regular Dirichlet form on $L^2(\mathbb{R}, m)$ with the effective intervals $\{(I_n, \mathbf{s}_n) : n \geq 1\}$ and assume $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$. Further let $\bar{\mathbf{s}}_n$ be the absolutely continuous part of \mathbf{s}_n and $\{\tilde{\mathcal{I}}_k : k \geq 1\}$ be the equivalence classes induced by the scale-connection with respect to $\{\bar{\mathbf{s}}_n : n \geq 1\}$. Then the closure $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$ of $C_c^\infty(\mathbb{R})$ in $(\mathcal{E}, \mathcal{F})$ can be expressed as (2.4) with the effective intervals $\{(\bar{\mathcal{I}}_k, \bar{\mathbf{s}}_k) : k \geq 1\}$, where $(\bar{\mathcal{I}}_k, \bar{\mathbf{s}}_k)$ is the C^∞ -merging of $\tilde{\mathcal{I}}_k$ for all k .*

Proof. Let $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$ be the Dirichlet form (2.4) with the effective intervals $\{(\bar{\mathcal{I}}_k, \bar{\mathbf{s}}_k) : k \geq 1\}$. It suffices to show $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}}) = (\mathcal{E}, \mathcal{F})$.

We first assert $C_c^\infty(\mathbb{R}) \subset \tilde{\mathcal{F}}$. In fact, since $g_n(x) > 0$ a.e. for any n and $\lambda_l(\bar{\mathcal{I}}_k) = 0$, it follows that $\bar{\mathbf{s}}_k$ is absolutely continuous on $\bar{\mathcal{I}}_k$ and $\bar{\mathbf{s}}'_k > 0$ a.e. on $\bar{\mathcal{I}}_k$. Thus the inverse function $\bar{\mathbf{t}}_k := \bar{\mathbf{s}}_k^{-1}$ is also absolutely continuous. Furthermore, for any $L > 0$,

$$\sum_k \int_{\bar{\mathcal{I}}_k \cap (-L, L)} \frac{1}{\bar{\mathbf{s}}'_k(x)} dx = \sum_n \int_{I_n \cap (-L, L)} \frac{1}{g_n(x)} dx.$$

Note that

$$\frac{1}{g_n(x)} dx = \left(\frac{dx}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n.$$

Hence we have

$$\sum_k \int_{\bar{\mathcal{I}}_k \cap (-L, L)} \frac{1}{\bar{\mathbf{s}}'_k(x)} dx = \sum_n \int_{I_n \cap (-L, L)} \left(\frac{dx}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n < \infty$$

by Lemma 3.3 and $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$. Applying Lemma 3.3 and Remark 3.4 (4) to $\{(\bar{\mathcal{I}}_k, \bar{\mathbf{s}}_k) : k \geq 1\}$, we have the inclusion $C_c^\infty(\mathbb{R}) \subset \tilde{\mathcal{F}}$. For any function $\varphi \in C_c^\infty(\mathbb{R})$, we have

$$\begin{aligned} \tilde{\mathcal{E}}(\varphi, \varphi) &= \frac{1}{2} \sum_k \int_{\bar{\mathcal{I}}_k} \varphi'(x)^2 \frac{1}{\bar{\mathbf{s}}'_k(x)} dx \\ &= \frac{1}{2} \sum_n \int_{I_n} \varphi'(x)^2 \left(\frac{dx}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n \\ &= \frac{1}{2} \sum_n \int_{I_n} \left(\frac{d\varphi}{d\mathbf{s}_n} \right)^2 d\mathbf{s}_n \\ &= \mathcal{E}(\varphi, \varphi) = \tilde{\mathcal{E}}(\varphi, \varphi). \end{aligned}$$

Finally, we need only to prove $C_c^\infty(\mathbb{R})$ is a special standard core of $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$. By the definition of the equivalence class, we can easily deduce that each $\bar{\mathcal{I}}_k$ is scale-isolated with respect to $\{\bar{\mathbf{s}}_k : k \geq 1\}$. Notice that $\bar{\mathbf{s}}_k$ is absolutely continuous. By applying Theorem 3.7 to $\{(\bar{\mathcal{I}}_k, \bar{\mathbf{s}}_k) : k \geq 1\}$, we conclude that $C_c^\infty(\mathbb{R})$ is a special standard core of $(\tilde{\mathcal{E}}, \tilde{\mathcal{F}})$. That completes the proof. \square

Remark 3.15. Consider the following example. Let $I_1 := (0, 1]$ and \mathbf{s}_1 be the scale function on I_1 given by (3.16). Further let $(\mathcal{E}, \mathcal{F})$ be the strongly local and regular Dirichlet form on $L^2(\mathbb{R})$ with only one effective interval (I_1, \mathbf{s}_1) , i.e.

$$\begin{aligned} \mathcal{F} &:= \left\{ u \in L^2(\mathbb{R}) : u|_{(0,1]} \in \mathcal{F}^{(\mathbf{s}_1)} \right\}, \\ \mathcal{E}(u, v) &:= \frac{1}{2} \int_0^1 \frac{du|_{(0,1]}}{d\mathbf{s}_1} \frac{dv|_{(0,1]}}{d\mathbf{s}_1} d\mathbf{s}_1, \quad u, v \in \mathcal{F}, \end{aligned}$$

where $\mathcal{F}^{(\mathbf{s}_1)}$ is given by (2.3) with $J = I_1, \mathbf{s} = \mathbf{s}_1$. Clearly, $C_c^\infty(\mathbb{R}) \subset \mathcal{F}$. Note that the absolutely continuous part of \mathbf{s}_1 is $\bar{\mathbf{s}}_1(x) = x$. Thus the C^∞ -merging of the unique equivalence

class is $\{(\bar{I}_1, \bar{s}_1)\}$ with $\bar{I}_1 = [0, 1]$ and $\bar{s}_1(x) = x$. It follows that the closure of $C_c^\infty(\mathbb{R})$ in $(\mathcal{E}, \mathcal{F})$ is

$$\begin{aligned}\bar{\mathcal{F}} &= \{u \in L^2(\mathbb{R}) : u|_{[0,1]} \in H^1([0,1])\}, \\ \bar{\mathcal{E}}(u, v) &= \frac{1}{2} \int_0^1 u'(x)v'(x)dx, \quad u, v \in \bar{\mathcal{F}}.\end{aligned}$$

The associated diffusion process on the effective interval $[0, 1]$ is the reflected Brownian motion.

In this example, 0 is a trap relative to $(\mathcal{E}, \mathcal{F})$ and thus $\{0\}$ is an exceptional set relative to $(\mathcal{E}, \mathcal{F})$. However, $\{0\}$ is not an exceptional set relative to $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$. Notice that $(\bar{\mathcal{E}}, \bar{\mathcal{F}})$ is a regular Dirichlet subspace of $(\mathcal{E}, \mathcal{F})$. As stated in [13, §2], if $(\mathcal{E}^1, \mathcal{F}^1)$ is a regular Dirichlet subspace of $(\mathcal{E}^2, \mathcal{F}^2)$ and denote their 1-capacities by Cap^1 and Cap^2 respectively, then $\text{Cap}^1(A) \geq \text{Cap}^2(A)$ for any appropriate set A . The above example tells us the strict inequality may hold. Intuitively speaking, a part of the state space with no sense for a Dirichlet form could be very important to its regular Dirichlet subspace (of course, not vice versa).

4. KILLING INSIDE

All the discussions above concern only the strongly local Dirichlet forms. That means the associated diffusion processes have no killing inside. This section is devoted to explain the following fact: the presence of killing inside has no influence for all the results above. The main tools we shall use are the resurrected transform and killing transform of the Dirichlet form, see [1, Theorems 5.2.17 and 5.1.6]. Note that any function in this section is taken as its quasi-continuous version.

We first consider a regular and local Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(I, m)$, where $I = \langle l, r \rangle$ and m is a fully supported Radon measure on I . Let k be the killing measure of $(\mathcal{E}, \mathcal{F})$ in its Beurling-Deny decomposition. Then k is a Radon smooth measure on I relative to $(\mathcal{E}, \mathcal{F})$. The resurrected Dirichlet form of $(\mathcal{E}, \mathcal{F})$ can be deduced as follows. Let \mathcal{F}_e be the extended Dirichlet space of $(\mathcal{E}, \mathcal{F})$ and define

$$\mathcal{E}^{\text{res}}(u, v) := \mathcal{E}(u, v) - \int_I u(x)v(x)k(dx), \quad u, v \in \mathcal{F}_e.$$

Clearly, $(\mathcal{E}^{\text{res}}, \mathcal{F}_e \cap L^2(I, m+k))$ is a regular Dirichlet form on $L^2(I, m+k)$. Denote the extended Dirichlet space of this Dirichlet form by $\mathcal{F}_e^{\text{res}}$ and set

$$\mathcal{F}^{\text{res}} := \mathcal{F}_e^{\text{res}} \cap L^2(I, m).$$

Then $(\mathcal{E}^{\text{res}}, \mathcal{F}^{\text{res}})$ is a regular and strongly local Dirichlet form on $L^2(I, m)$ sharing the same set of quasi notions with $(\mathcal{E}, \mathcal{F})$ (Cf. [1, Theorem 5.2.17]). It is called the resurrected Dirichlet form of $(\mathcal{E}, \mathcal{F})$. Furthermore, we have

$$(4.1) \quad \mathcal{F}^{\text{res}} \cap L^2(I, k) = \mathcal{F}_e^{\text{res}} \cap L^2(I, m+k) = \mathcal{F}_e \cap L^2(I, m+k) = \mathcal{F}$$

and

$$(4.2) \quad \mathcal{E}(u, v) = \mathcal{E}^{\text{res}}(u, v) + \int_I u(x)v(x)k(dx), \quad u, v \in \mathcal{F}^{\text{res}} \cap L^2(I, k) = \mathcal{F}.$$

Thus on the contrary, $(\mathcal{E}, \mathcal{F})$ is the killed Dirichlet form of $(\mathcal{E}^{\text{res}}, \mathcal{F}^{\text{res}})$ induced by k . Since Theorem 2.1 gives an expression of $(\mathcal{E}^{\text{res}}, \mathcal{F}^{\text{res}})$, we can obtain the following characterization of $(\mathcal{E}, \mathcal{F})$.

Theorem 4.1. *Let $I = \langle l, r \rangle$ be an interval and m a fully supported Radon measure on I . Then $(\mathcal{E}, \mathcal{F})$ is a regular and local Dirichlet form on $L^2(I, m)$ if and only if there exist a set of at most countable disjoint intervals $\{I_n = \langle a_n, b_n \rangle : I_n \subset I, n \geq 1\}$, a scale function $s_n \in \mathbf{S}_\infty(I_n)$ for each $n \geq 1$ and a Radon measure k on I with $k \ll m$ on $I \setminus \cup_{n \geq 1} I_n$, such that*

$$(4.3) \quad \begin{aligned}\mathcal{F} &= \left\{ u \in L^2(I, m+k) : u|_{I_n} \in \mathcal{F}^{(s_n)}, \sum_{n \geq 1} \mathcal{E}^{(s_n)}(u|_{I_n}, u|_{I_n}) < \infty \right\}, \\ \mathcal{E}(u, v) &= \sum_{n \geq 1} \mathcal{E}^{(s_n)}(u|_{I_n}, v|_{I_n}) + \int_I u(x)v(x)k(dx), \quad u, v \in \mathcal{F},\end{aligned}$$

where for each $n \geq 1$, $(\mathcal{E}^{(\mathbf{s}_n)}, \mathcal{F}^{(\mathbf{s}_n)})$ is given by (2.3) with the scale function \mathbf{s}_n . Moreover, the intervals $\{I_n : n \geq 1\}$ and scale functions $\{\mathbf{s}_n : n \geq 1\}$ are uniquely determined, if the difference of order is ignored.

Proof. For the necessity, let k be the killing measure of $(\mathcal{E}, \mathcal{F})$ and consider the resurrected Dirichlet form $(\mathcal{E}^{\text{res}}, \mathcal{F}^{\text{res}})$ of $(\mathcal{E}, \mathcal{F})$. Since $(\mathcal{E}^{\text{res}}, \mathcal{F}^{\text{res}})$ is a regular and strongly local Dirichlet form on $L^2(I, m)$, it follows that $(\mathcal{E}^{\text{res}}, \mathcal{F}^{\text{res}})$ is expressed as (2.4) with a class of effective intervals $\{(I_n, \mathbf{s}_n) : n \geq 1\}$. Since $(\mathcal{E}, \mathcal{F})$ and $(\mathcal{E}^{\text{res}}, \mathcal{F}^{\text{res}})$ share the same set of quasi notions, and the subset of $I \setminus \cup_{n \geq 1} I_n$ with zero m -measure is exceptional relative to $(\mathcal{E}^{\text{res}}, \mathcal{F}^{\text{res}})$, we can conclude $k \ll m$ on $I \setminus \cup_{n \geq 1} I_n$. Finally, (4.1) and (4.2) implies the expression (4.3) of $(\mathcal{E}, \mathcal{F})$.

For the sufficiency, we consider the regular Dirichlet form (2.4) on $L^2(I, m)$. Note that the exceptional subset of I_n relative to (2.4) must be the empty set. Then $k \ll m$ on $I \setminus \cup_{n \geq 1} I_n$ implies k is a Radon smooth measure relative to (2.4) and thus (4.3) is the killed Dirichlet form of (2.4) induced by k . Therefore, (4.3) is a regular and local Dirichlet form on $L^2(I, m)$. That completes the proof. \square

We end this section with some explanations for the extensions of Theorems 3.7 and 3.14 to a local Dirichlet form. As proved above, (2.4) is the resurrected Dirichlet form of (4.3) and (4.3) is the killed Dirichlet form of (2.4) induced by k . From [1, Theorems 5.1.6 and 5.2.17], we know that a class \mathcal{C} is a special standard core of (2.4) if and only if \mathcal{C} is a special standard core of (4.3). Therefore, nothing else need to be changed if the Dirichlet forms in Theorems 3.7 and 3.14 are only assumed to enjoy the local property.

ACKNOWLEDGEMENT

Example 2.10 was inspired by the discussions with Professor Zhi-Ming Ma, and the first named author would like to thank his helpful advice.

REFERENCES

- [1] Chen, Z.-Q., Fukushima, M.: Symmetric Markov processes, time change, and boundary theory. Princeton University Press, Princeton, NJ (2012).
- [2] Fang, X., Fukushima, M., Ying, J.: On regular Dirichlet subspaces of $H^1(I)$ and associated linear diffusions. Osaka J. Math. 42, 27-41 (2005).
- [3] Fang, X., He, P., Ying, J.: Dirichlet forms associated with linear diffusions. Chin. Ann. Math. Ser. B. 31, 507-518 (2010).
- [4] Fitzsimmons, P. J., Li, L.: Class of smooth functions in Dirichlet spaces. arXiv: 1611.06778.
- [5] Folland, G.B.: Real analysis. John Wiley & Sons, Inc., New York (1999).
- [6] Fukushima, M.: From one dimensional diffusions to symmetric Markov processes. Stochastic Process. Appl. 120 (2010), no. 5, 590-604.
- [7] Fukushima, M.: On general boundary conditions for one-dimensional diffusions with symmetry. J. Math. Soc. Japan. 66, 289-316 (2014).
- [8] Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet forms and symmetric Markov processes. Walter de Gruyter & Co., Berlin (2011).
- [9] Hamza, M.: Détermination des formes de Dirichlet sur \mathbb{R}^n , Thèse 3e cycle, Orsay, 1975.
- [10] Itô, K.: Essentials of stochastic processes. American Mathematical Society, Providence, RI (2006).
- [11] Itô, K., McKean, H.P., Jr: Diffusion processes and their sample paths. Springer-Verlag, Berlin-New York (1974).
- [12] Li, L., Ying, J.: Regular subspaces of Dirichlet forms. In: Festschrift Masatoshi Fukushima. pp. 397-420. World Sci. Publ., Hackensack, NJ (2015).
- [13] Li, L., Ying, J.: Regular Dirichlet extensions of one-dimensional Brownian motion, (2016). arXiv: 1606.00630.
- [14] Ma, Z.M., Röckner, M.: Introduction to the theory of (nonsymmetric) Dirichlet forms. Springer-Verlag, Berlin, Heidelberg (1992).
- [15] Rogers, L.C.G., Williams, D.: Diffusions, Markov processes, and martingales. Vol. 2. John Wiley & Sons, Inc., New York (1987).

INSTITUTE OF APPLIED MATHEMATICS, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES, BEIJING 100190, CHINA.

E-mail address: liping_li@amss.ac.cn

SCHOOL OF MATHEMATICAL SCIENCES, FUDAN UNIVERSITY, SHANGHAI 200433, CHINA.

E-mail address: jgying@fudan.edu.cn